Duality of stochasticity and natural selection: a cybernetic evolution theory

*Peer review status:*
No

**Corresponding Author:**
Prof. Kurt Heininger,
Professor, Department of Neurology, Heinrich Heine University Duesseldorf - Germany

**Submitting Author:**
Prof. Kurt Heininger,
Professor, Department of Neurology, Heinrich Heine University Duesseldorf - Germany

**Article ID:** WMC004796
**Article Type:** Original Articles
**Submitted on:** 22-Feb-2015, 09:03:27 PM GMT  **Published on:** 23-Feb-2015, 07:42:40 AM GMT
**Article URL:** http://www.webmedcentral.com/article_view/4796
**Subject Categories:** ECOLOGY
**Keywords:** Stochasticity, natural selection, cybernetics, bet-hedging, multilevel selection, Law of Requisite Variety, mean geometric fitness

**How to cite the article:** Heininger K. Duality of stochasticity and natural selection: a cybernetic evolution theory. WebmedCentral ECOLOGY 2015;6(2):WMC004796

**Copyright:** This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC-BY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Source(s) of Funding:**
No source of funding

**Competing Interests:**
No competing interests
Duality of stochasticity and natural selection: a cybernetic evolution theory

Author(s): Heininger K

Abstract

Orthodox Darwinism assumes that environments are stable. There is an important difference between breeding (Darwin’s role model of evolution) and evolution itself: while in breeding the final goal is preset and constant, adaptation to varying biotic and abiotic environmental conditions is a moving target and selection can be highly fluctuating. Evolution is a cybernetic process whose Black Box can be understood as learning automaton with separate input and output channels. Cybernetics requires a closed signal loop: action by the system causes some change in its environment and that change is fed to the system via information (feedback) that enables the system to change its behavior. The input signal is given by a complex biotic and abiotic environment. Natural selection is the output/outcome of the learning automaton.

Environments are stochastic. Particularly, density- and frequency-dependent coevolutionary interactions generate chaotic and unpredictable dynamics. Stochastic environments coerce organisms into risky lotteries. Chance favors the prepared. The ‘Law of Requisite Variety’ holds that cybernetic systems must have internal variety that matches their external variety so that they can self-organize to fight variation with variation. Both conservative and diversifying bet-hedging are the risk-avoiding and spreading insurance strategies in response to environmental uncertainty. The bet-hedging strategy tries to cover all bases in an often unpredictable environment where it does not make sense to “put all eggs into one basket”. In this sense, variation is the bad/worst-case insurance strategy of risk-aversive individuals. Variation is pervasive at every level of biological organization and is created by a multitude of processes: mutagenesis, epimutation, recombination, transposon mobility, repeat instability, gene expression noise, cellular network dynamics, physiology, phenotypic plasticity, behavior, and life history strategy. Importantly, variation is created condition-dependently, when variation is most needed – in organisms under stress. The bet-hedging strategy also manifests in a multitude of life history patterns: turnover of generations, reproductive prudence, iteroparity, polyandry, and sexual reproduction.

Cybernetic systems are complex systems. Complexity is conceived as a system’s potential to assume a large number of states, i.e., variety. Complex systems have both stochastic and deterministic properties and, in fact, generate order from chaos. Non-linearity, criticality, self-organization, emergent properties, scaling, hierarchy and evolvability are features of complex systems. Emergent properties are features of a complex system that are not present at the lower level but arise unexpectedly from interactions among the system’s components. Only within an intermediate level of stochastic variation, somewhere between determined rigidity and literal chaos, local interactions can give rise to complexity. Stochastic environments change the rules of evolution. Lotteries cannot be played and insurance strategies not employed with single individuals. These are emergent population-level processes that exert population-level selection pressures generating variation and diversity at all levels of biological organization. Together with frequency and density-dependent selection, lottery- and insurance-dependent selection act on population-level traits.

The duality of stochasticity and selection is the organizing principle of evolution. Both are interdependent. The feedback between output and input signals inextricably intertwines both stochasticity and natural selection, and the individual- and population-levels of selection. Sexual reproduction with its generation of pre-selected variation is the paradigmatic bet-hedging enterprise and its evolutionary success is the selective signature of stochastic environments. Sexual reproduction is the proof of concept that (epi)genetic variation is no accidental occurrence but a highly regulated process and environmental stochasticity is its evolutionary “raison d’être”. Evolutionary biology is plagued by a multitude of controversies (e.g. concerning the level of selection issue and sociobiology. Almost miraculously, these controversies can be resolved by the cybernetic model of evolution and its implications.

Table of contents

Abstract
Table of contents

1. Introduction
2. Darwin's role model of evolution
3. The cybernetics of evolution
4. The improbability of evolution
5. The cybernetic learning automaton
6. Black Box theory
   6.1 Input and output of the Back Box
7. Evolutionary outcomes are probabilistic. But why?
8. Chance and necessity
   8.1 Chance
   8.1.1 Mutations
   8.1.2 Genetic drift
   8.2 Necessity
9. Gedankenexperiment: evolution in stable environments
10. The evolutionary signature of stochastic environments
   10.1 Environmental stochasticity
   10.2 Demographic stochasticity
   11.1 Molecular biological bet-hedging
   11.1.1 Gene expression noise
   11.1.2 Epigenesis
   11.1.3 Protein promiscuity
   11.1.4 Energy-Ca2+-redox triangle
   11.2 Phenotypic and behavioral bet-hedging
   11.2.1 Canalization and phenotypic plasticity: two sides of the same coin
   11.3 Stress and bet-hedging
12. The gambles of life
   12.1 Lottery and insurance: responses to uncertainty and risk
   12.2 "Decisions" under uncertainty: utility/fitness optimization
13. Evolution is far-sighted
14. Complexity and self-organization: chaos and order
   14.1 Complexity
   14.2 Fractals and 1/f noises
   14.3 Self-organization
   14.3.1 Self-organized criticality
15. Stochasticity and multilevel selection
   15.1 Community selection as an emergent behavior of complex systems
   15.2 Fitness as transgenerational propensity
   15.3 Reproductive fitness in stochastic environments
   15.3.1 Reproductive prudence
16. Life history phenotypes of bet-hedging
   16.1 Turnover of generations: bet-hedging in time?
   16.2 Iteroparity
   16.3 Polyandry
   16.4 Sexual reproduction
17. Stochasticity and selection: duality in evolution
   17.1 The creative conflict between stochastic indeterminism and selective determinism
   17.1.1 Playing dice with controlled odds
18. The blending of ecology and evolution
19. Cutting the Gordian knot of controversies
20. Abbreviations
21. References

1. Introduction

Although biology has been the theater of numerous controversies, the view that biological processes must be deterministic has almost never been put into question. Since Ancient times with Aristotle and his finalist conception of living beings, since Descartes and his mechanistic theory of life in the 17th century, since Claude Bernard and his physico-chemical description of physiology in the 19th century, until the computational metaphor of a genetic program that was put forth by Monod and Jacob in the 20th century, biology has always been dominated by deterministic theories. One may even say that determinism has always appeared as being in the deep nature of life, and therefore dominant in the explanation of the latter. Gandrillon et al. (2012)

Evolutionary forces are often divided into two sorts: stochastic and deterministic (Wright, 1955). To date, there is a general agreement that ecological and evolutionary outcomes and the tools that are used to describe them are probabilistic and statistical (Beatty, 1984; Millstein, 1997, 2003; Graves et al., 1999; Glymour, 2001; Shanahan, 2003; Rosenberg, 2004; Colyvan, 2005). However, it has been a contentious issue whether evolution is deterministic or indeterministic (Rosenberg, 1994; Horan, 1994; Brandon & Carson, 1996; Graves et al., 1999; Stamos, 2001; Weber, 2001; Shanahan, 2003). Glymour (2001)
concluded that "any complete and correct evolutionary theory must be probabilistic", a statement that has not been questioned by the advocates of a more deterministic perspective (Rosenberg, 2001). However, the deterministic perspective (e.g. Graves et al., 1999; Rosenberg, 2001) reflects the dictum of Laplace (1825) that randomness is only a measure of our "ignorance of the different causes involved in the production of events." "The world, it is said, might often look haphazard, but only because we do not know the inevitable workings of its inner springs" (Hacking, 1990, p. 1).

Contingency may be defined as the outcome of a particular set of concomitant effects that apply in a particular space-time situation and thus determine the outcome of a given event. In most of the epistemological literature, this word has aptly taken the place of the term 'chance' or 'random event', and in fact, it has a different texture. For example, a car accident can be seen as a chance event, but indeed it is due to the concomitance of many independent factors, like the car speed, the road conditions, the state of the tyres, the alcohol consumption of the driver, etc. These factors all sum together to give the final result, seen as a chance event. The same can be said for a stock-market crash, or the stormy weather of a particular summer day. Interestingly, each of these independent factors can actually be seen per se as a deterministic event, e.g. the bad state of the car tyres determines per se a car sliding off at a curve. The fact, however, that there are so many of these factors, and each with an unknown statistical weight, renders the complete accident unpredictable: a chance event. Change the contingent conditions (perhaps only one of them) and the final result would be quite different: it may happen one week later, or with another driver, or never. As Gould states about biological evolution (Gould, 1989), "run the tape again, and the first step from prokaryotic to eucariotic cell may take 12 billion years instead of two. . . ", implying that the onset of the determinism-natural selection duality in evolution that was first presented under the impression that sexual reproduction is a sophisticated bet-hedging enterprise in response to environmental stochasticity (Heininger, 2013). I will argue that from a cybernetic perspective there is compelling evidence for a dualistic creative conflict between stochasticity and selection in evolution. The dictionary definition of stochastic is "(1) Relating to, or characterized by conjecture; conjectural. (2) Involving or containing a random variable or variables: stochastic calculus. (3) Involving chance or probability: a stochastic stimulation. (4) adj: (statistics) being or having a random variable; "a stochastic variable"; "stochastic processes." The ambiguous implication of chance, randomness and probability reflects the uncertainty and unpredictability of abiotic and biotic systems and the probabilistic character of evolution (Weiss & Buchanan, 2011).

2. Darwin

The Modern Synthesis built on Darwin's two major realizations: (i) that all living organisms are related to one another by common descent; (ii) that a primary
explanation for the pattern of diversity of life—and especially for the obvious “fit” of organisms to their environments—is the process that he called natural selection. He recognized the importance of variation for the action of selection (1859, chapters I, II and V). However, he had no idea how this variation arose: “our ignorance of the laws of variation is profound” (1859, p. 149). Yet, understanding the “laws of variation” should have a key role in our understanding of evolution. At the beginning of the 20th century the foundation of the modern theory of evolution posited that evolution is the result of the interplay between two antagonistic mechanisms: natural selection and sources of genetic variation (Campos & Wahl, 2010). In Darwin’s tradition, the Modern Synthesis understood selection as the only driving force in evolution (Mayr, 1980). Genetic variation was considered the result of accidental mutations. But Levinton (1988, p. 494) stated: “Evolutionary biologists have been mainly concerned with the fate of variability in populations, not the generation of variability. ... This could stem from the dominance of population genetic thinking, or it may be due to a general ignorance of the mechanistic connections between the genes and the phenotype. Whatever the reason, the time has come to reemphasize the study of the origin of variation.” During the 27 years since this statement our knowledge of the proximal, molecular biological, generation of variation has been expanded tremendously (see chapter 11.1) but little progress has been made in our understanding of the ultimate evolutionary origin of variation.

Darwin’s strongest evidence for the power of natural selection was by analogy with the dramatic success of artificial selection (Darwin, 1859, Chapter 1) and studies since Darwin’s time have confirmed his view. What is remarkable is that almost all traits respond to selection, and that selection on large sexual populations causes a sustained response over many generations (Barton & Turelli, 1989; Falconer & Mackay, 1995). But there is an important difference between artificial selection and evolution. In breeding, artificial selection has the goal to improve a certain predefined trait, e.g. oil content in maize (Laurie et al., 2004), milk quantity and quality in dairy cattle breeding (Miglior et al., 2013), a certain morphological trait in pigeons, or, in the laboratory, flight speed in Drosophila (Weber, 1996). The target is pre-defined by the breeder (figure 1B). Importantly, breeding is an iterative process (Hill & Caballero, 1992; Williams & Lenton, 2007) in which the ultimate goal is reached after many generations, but the setting of the ultimate goal and thus the direction of selection remain constant. Here, variation is an often unwanted noise, at least when it does not serve the ultimate target of selection. The breeder has at least two functions: (s)he determines the goal of the breeding operation and selects the individuals for the next round of breeding. In evolution, however, the direction and selective regime are established by the organism’s stochastic, often unpredictable, environment. The evolutionary dynamics consist of a fitness-dominated, directed part caused by selection and a neutral, undirected part due to fluctuations (Frey, 2010). Thus, adaptation to varying biotic and abiotic environmental conditions is determined by a moving target and selection can be highly fluctuating (figure 1C; Siepielski et al., 2009; Bell, 2010).

3. The cybernetics of evolution

There is simply no denying the breathtaking brilliance of the designs to be found in nature. Time and again, biologists baffled by some apparently futile or maladroit bit of bad design in nature have eventually come to see that they have underestimated the ingenuity, the sheer brilliance, the depth of insight to be discovered in one of Mother Nature's creations. Francis Crick......baptized this trend in the name of his colleague Leslie Orgel, speaking of what he calls “Orgel’s Second Rule: Evolution is cleverer than you are.”

Daniel Dennett, Darwin’s Dangerous Idea, 1995

The term cybernetics stems from the Greek κυβερνητης (kybernetes = steersman, governor, pilot, or rudder). Cybernetics had a crucial influence on the birth of various modern sciences: control theory, computer science, information theory, automata theory, artificial intelligence and artificial neural networks, cognitive science, computer modeling and simulation science, dynamical systems, and artificial life. Many concepts central to these fields, such as complexity, self-organization, self-reproduction, autonomy, networks, connectionism, and adaptation, were first explored by cyberneticians during the 1940’s and 1950’s. Examples include von Neumann’s computer architectures, game theory, and cellular automata; Ashby’s and von Foerster’s analysis of self-organization; Bratenberg’s autonomous robots; and McCulloch’s artificial neural nets, perceptrons, and classifiers (Heylighen & Joslyn, 2001). Cybernetics is the science of control systems; or, to expand it into Norbert Wiener’s own words: “the science of control and communication in the animal and the machine” (Wiener, 1948, Ashby, 1956). According to Beer (1959), there are three main characteristics of a cybernetic system: extreme
complexity, probabilism, and self-regulation. Cybernetics is about how to cope with the challenge of ubiquitous complexity (see also chapter 14). Cybernetic systems are characterized by input and output variables and it is essential to distinguish the one from the other (Ashby, 1956). The control of a system requires getting information from the output back to the input of a system and this is called feedback. Cybernetics requires a closed signal loop: action by the system causes some change in its environment and that change is fed to the system via information (feedback) that enables the system to change its behavior. Cybernetic systems are systems with feedback. They are a special class of cause-and-effect (input-output) systems (figure 1A). Patten and Odum (1981) offered a minimalist definition that distinguished cybernetic systems from non-cybernetic systems by the presence of feedback control; in cybernetic systems, “input is determined, at least in part, by output”. Very small feedbacks may exert very large effects (Patten & Odum, 1981).

All life is cybernetic (Korzeniewski, 2001, 2005; De Silva & Uchiyama, 2007; Nurse, 2008; Abel, 2012). All life depends upon linear digital prescriptive information and cybernetic programming (Abel, 2012). Genetic cybernetics even inspired Turing’s, von Neumann’s, and Wiener’s development of computer science (Turing, 1936; von Neumann, 1950, 1956; von Neumann et al., 1987, 2000; Wiener, 1948, 1961). Evolution is a cybernetic process (e.g. Ashby, 1954; 1956; Schmalhausen, 1960; Waddington, 1961; Corning, 2005; Scott, 2010). In his 1858 essay, A.R. Wallace referred to the evolutionary principle “as exactly like that of the centrifugal governor of the steam engine, which checks and corrects any irregularities almost before they become evident....”.

For Gregory Bateson (1972, p. 435) “the result will be...a self-corrective system. Wallace, in fact, proposed the first cybernetic model.” However, the first account of how a phenotypic change induced by a change in the environment could lead to a change in the inherited genome was provided by Spalding (1837). Spalding’s driver of evolution comprised a sequence of learning followed by differential survival of those individuals that expressed the phenotype more fit, even granting it billions of years? In the view of mathematicians, the ratio of the number of functional genes and proteins, on the one hand, to the enormous number of possible sequences corresponding to a gene or protein of a given length, on the other, seemed so small as to preclude the origin of genetic information by a random mutational search. A functional protein one hundred amino acids in length represents an extremely unlikely occurrence. There are roughly $10^{100}$ possible amino acid sequences of this length, if one considers only the 20 protein-forming acids as possibilities. In human codes, M. P. Schützenberger (1967) argued, randomness is never the friend of function, much less of progress. When we make changes randomly to computer programs, “we find that we have no chance (i.e. less than 1/10^{100}) even to see what the modified program would compute: it just jams.” Bak (1993, cited in Fernández et al., 1998) described the difficulty: “If, for the sake of argument, we imagine the outer world frozen (for a while) and try to construct from scratch an equally fit species by recourse to engineering techniques rather than by evolution, we will be forced to accept that eons are needed. By starting at a random configuration one certainly will reach a wrong and much less fit maximum. It would be necessary to systematically go through all configurations, involving

Organ’s Second Rule: “Evolution is cleverer than you are.”

4. The improbability of evolution

Mathematical models that mimic biological evolutionary processes have revealed that the traditional view of Darwinian evolution, according to which the most fit of random mutants are selected, faces a major problem (Eden, 1967; Schützenberger, 1967; Bak et al., 1987, 1988; Bak, 1993, 1996; Kauffman, 1995 p. 155ff; Fernández et al., 1998): It is much too slow to account for real evolution. In 1966, mathematicians, physicists and engineers met in Philadelphia (Moorhead & Kaplan, 1967). The mathematicians argued that neo-Darwinism faced a formidable combinatorial problem. Murray Eden illustrated the issue with reference to an imaginary library evolving by random changes to a single phrase: “Begin with a meaningful phrase, retype it with a few mistakes, make it longer by adding letters, and rearrange subsequences in the string of letters; then examine the result to see if the new phrase is meaningful. Repeat until the library is complete” (Eden, 1967). Would such an exercise have a realistic chance of succeeding, even granting it billions of years? In the view of mathematicians, the ratio of the number of functional genes and proteins, on the one hand, to the enormous number of possible sequences corresponding to a gene or protein of a given length, on the other, seemed so small as to preclude the origin of genetic information by a random mutational search. A functional protein one hundred amino acids in length represents an extremely unlikely occurrence. There are roughly $10^{100}$ possible amino acid sequences of this length, if one considers only the 20 protein-forming acids as possibilities. In human codes, M. P. Schützenberger (1967) argued, randomness is never the friend of function, much less of progress. When we make changes randomly to computer programs, “we find that we have no chance (i.e. less than 1/10^{100}) even to see what the modified program would compute: it just jams.” Bak (1993, cited in Fernández et al., 1998) described the difficulty: “If, for the sake of argument, we imagine the outer world frozen (for a while) and try to construct from scratch an equally fit species by recourse to engineering techniques rather than by evolution, we will be forced to accept that eons are needed. By starting at a random configuration one certainly will reach a wrong and much less fit maximum. It would be necessary to systematically go through all configurations, involving
Learning is a process of acquiring information, storing it in memory, and using it to modify future behaviors. A learning system is characterized by its ability to improve its behavior with time, in some sense tending towards an ultimate goal (Narendra & Thathachar, 1974). The evolution of learning is a paradigm case of the dual action of environmental stochasticity and natural selection. The ability to learn is a behavioral capacity whose evolution is usually explained through the action of natural selection (e.g. Staddon, 1983; Marler & Terrace, 1984; Bolles & Beecher, 1988; Davey, 1989; Miller & Todd, 1990). However, a vital component of the learning process is also the environment. If the environments were relatively static, there might be little need for learning to evolve. Since some cost is associated with learning (Dukas & Duan, 2000; Mery & Kawecki, 2003, 2004; Burger et al., 2008), in an absolutely fixed environment a genetically fixed pattern of behavior should evolve (“absolute fixity argument”). But if the environment is diverse and unpredictable, innate environment-specific mechanisms are of little use. Unpredictable or variable environments favor the evolution of cognition and learning (Bergman & Feldman, 1995; Richerson & Boyd, 2000; Godfrey-Smith, 2002; Mery & Kawecki, 2002; Brown et al., 2003; Kerr & Feldman, 2003; Heller, 2004; Kotrschal & Taborsky, 2010; Richardson, 2012; Clarin et al., 2013; Tebbich & Teschke, 2014) and cognition/learning is thought to enable organisms to deal with environmental heterogeneity (Godfrey-Smith, 2002; Richardson, 2012). “The function of cognition is to enable the agent to deal with environmental complexity” (Godfrey-Smith, 1996). Learning is an important pathway to flexibility as it allows animals to adjust their behavior to environmental changes. On the other hand, in an absolutely unpredictable environment, where the past and present states of the environment offer no information about the future there is nothing to learn and there is again no driving force for a learning capability to evolve (Bergman & Feldman, 1995). Thus, learning should only be adaptive, if learning rates are sufficiently higher than the rates of environmental change (Dukas, 1998) and should therefore vary with environmental stability and predictability. Similarly, Stephens (1991) argued that the pattern of predictability in relation to an individual’s life history could determine the evolutionary advantage of learning. Within these framework conditions, a stochastic environment encourages the evolution of learning (Levins, 1968; Johnston, 1982; Chalmers, 1990; Stephens, 1991; Bergman & Feldman, 1995; Krakauer & Rodríguez-Gironés, 1995; Groß et al., 2008; Eliassen et al., 2009). Bergman and Feldman (1995) viewed learning as the ability to construct a representation of the environment and, by proper use of the representation, to predict future states of the environment. This requires some regularity in the environmental signals and the capacity to capture this regularity. Learning is believed to be adaptive because under a wide range of conditions it allows the learner to generate predictions about its environment, and hence to make better decisions, than by using innate knowledge alone (Johnston, 1982; Stephens, 1991; Bergman & Feldman, 1995). Environmental fluctuations early in life are known to enhance the behavioral flexibility of animals with regard to predator avoidance strategies (Braithwaite & Salvanes, 2005; Salvanes et al., 2007), feeding performance (Braithwaite & Salvanes, 2005) and social behavior (Salvanes & Braithwaite, 2005; Salvanes et al., 2007). A possible explanation for these behavioral effects is that variable environments evoke repeated neural stimulations resulting in faster and better learning (Braithwaite & Salvanes, 2005). Several studies showed that neural stimulation over longer periods by exposing animals to enriched environments (e.g., Kempermann et al., 1997; Brown et al., 2003) can enhance brain development (Bredy et al., 2004; Botero et al., 2009) for example through an increased synaptic density (Bredy et al., 2003), and can lead to improved learning abilities and memory capacity (Bredy et al., 2003). For example, the learning abilities of fishes increased in response to experimental variation of environmental quality during ontogeny (Kotrschal & Taborsky, 2010).

Learning automata are adaptive decision-making devices operating on unknown random environments (Narendra & Thathachar, 1974; 1989). The automaton
updates its action probabilities in accordance with the inputs received from the environment so as to improve its performance in some specified sense (Narendra & Thathachar, 1974; 1989). The basic operation carried out by a learning automaton is the updating of the action probabilities on the basis of the responses of the environment. The learning automaton has a finite set of actions and each action has a certain probability (unknown to the automaton) of getting rewarded by the controlled system, which is considered as environment of the automaton. The aim is to learn to choose the optimal action (i.e. the action with the highest probability of being rewarded) through repeated interaction on the system. If the learning algorithm is chosen properly, then the iterative process of interacting on the system can be made to result in the selection of the optimal action (Zeng et al., 2000). The learning model that is closest to the evolutionary approach is “reinforcement learning” based on the insight that successful strategies will be reinforced and used more frequently. Reinforcement learning has been successfully applied for solving problems involving decision making under uncertainty (Narendra & Thathachar, 1989; Barto et al., 1983; Zikidis & Vasilakos, 1996; Zeng et al., 2000; Thathachar & Sastry, 2002). (When speaking of ‘decisions’, use of the term is in an evolutionary sense, not implying any conscious rationalization on the part of individual organisms.) In general, a reinforcement learning algorithm conducts a stochastic search of the output space, using only an approximate indication of the “correctness” (reward) of the output value it produced in every iteration. Based on this indication, a reinforcement learning algorithm generates, in each iteration, an error signal giving the difference between the actual and correct response and the adaptive element uses this error signal to update its parameters (Zeng et al., 2000).

6. Black Box theory

“In our daily lives we are confronted at every turn with systems whose internal mechanisms are not fully open to inspection, and which must be treated by the methods appropriate to the Black Box” (Ashby, 1956, p. 86). A Black Box theory treats its object or subject matter as if it were a system devoid of internal structure; it focuses on the system’s behavior and handles the system as a single unit (Bunge, 1967 p. 509). The cybernetic Black Box theory deals with incomplete knowledge about causal mechanisms but deduces knowledge about the system’s properties from the relations between the input- and the corresponding output-characteristics (Ashby, 1956 chapter 6). A complex system usually cannot be studied by decomposing the system into its constituent subsystems, but rather by measuring certain signals generated by the system and analyzing the signals to gain insights into the behavior of the system (Gao et al., 2007). Since it is often difficult to predict the behavior of a complex system, Simon (1981) recommends vicarious system experimentation through simulation, pointing out that this technique may even create new knowledge about system behavior. He is especially keen to demonstrate that system behavior can be predicted even in ignorance of (or with a minimal knowledge of) the system’s structure. In connection with this, he speaks in favor of Black Box theories (Simon, 1981, p. 20): “We knew a great deal about the gross physical and chemical behavior of matter before we had a knowledge of molecules, a great deal about molecular chemistry before we had an atomic theory, and a great deal about atoms before we had any theory of elementary particles—if indeed we have such a theory today. This skyhook-skyscraper construction of science from the roof down to the yet unconstructed foundations was possible because the behavior of the system at each level depended on only a very approximate, simplified, abstracted characterization of the system at the level next beneath.” Simon also refers to John von Neumann’s research in computer reliability and the problem of organizing a system in such a way that as a whole, it becomes relatively reliable in spite of the possible unreliability of its components (Mattessich, 1982).

6.1 Input and output of the Black Box

Darwin was vague in the meaning of his new concept of “Natural Selection,” using it interchangeably as one of the causes for evolutionary change and as the final outcome (= evolutionary change). But his clearest definition of natural selection (Darwin, 1859 p. 61: “I have called this principle, by which each slight variation, if useful, is preserved, by the term of Natural Selection, in order to mark its relation to man’s power of selection.”) is an outcome definition, not that of a cause (Bock, 2003). First, natural selection is a metaphor, an umbrella term that serves to label and characterize a vast array of specific factors with survival consequences. The generally accepted modern definition of natural selection is that it is an outcome (Fisher, 1930; Endler, 1986; Bock, 2003, 2010; Reese, 2005), and is: “nonrandom (differential) reproduction of genotypes” (e.g., Ehrlich & Holm, 1963, p. 326); or “nonrandom differential survival or reproduction of classes of phenotypically different
entities.” (Futuyma, 1986, p. 555). Natural selection is used by most biologists (e.g., Dobzhansky, 1959; Lerner, 1959) quite interchangeably as a cause, a process and an outcome resulting in massive confusion (Endler, 1986; Bock, 2010; MacColl, 2011). However, since evolution is a continuous iterative process where the resultant population of the previous contest becomes the input population to the next contest, Darwin’s and his followers’ ambiguity was not unfounded.

In a cybernetic system, natural selection corresponds to an output. Surprisingly, the input into the process of evolution has been less rigorously defined. Ross Ashby (1956, p. 46) defined the input of cybernetic systems as follows: “With an electrical system, the input is usually obvious and restricted to a few terminals. In biological systems, however, the number of parameters is commonly very large and the whole set of them is by no means obvious. It is, in fact, co-extensive with the set of ‘all variables whose change directly affects the organism’. The parameters thus include the conditions in which the organism lives. In the chapters that follow [in Ashby’s book, _KH_], the reader must therefore be prepared to interpret the word “input” to mean either the few parameters appropriate to a simple mechanism or the many parameters appropriate to the free-living organism in a complex environment.” Environmental information falls into two categories: signals and environmental factors. The signals are deterministic, that is, once received the consequences are inevitable. The environmental factors are stochastic, that is they generate randomness (Skår & Coveney, 2003). Complex environments have a high degree of uncertainty/stochasticity (Yoshimura & Clark, 1991; Grant & Grant, 2002; Lenormand et al., 2009) or capriciousness (Lewontin, 1966). Uncertainty in environments is a function of (1) degrees of freedom (generally taken as the most basic definition of complexity [Gell-Mann, 1994]); (2) the possible nonlinearity of each variable comprising each degree of freedom, and (3) the possibility that each may change (McKelvey, 2004a).

Stochastic automata operating in an unknown random environment have been proposed as models of learning (Narendra & Thathachar, 1974). These automata update their action probabilities in accordance with the inputs received from the environment and can improve their own performance during operation. Developments in stochastic control theory took into account uncertainties that might be present in the process; stochastic control was effected by assuming that the probabilistic characteristics of the uncertainties are known. Frequently, however, the uncertainties are of a higher order, and even the probabilistic characteristics such as the distribution functions may not be completely known. It is then necessary to make observations on the process as it is in operation and gain further knowledge of the process. A distinctive feature of such problems is that there is little _a priori_ information, and additional information is to be acquired on line. Narendra and Thathachar (1974) illustrated the automaton approach in an example featuring a student and a probabilistic teacher. “A question is posed to the student and a finite set of alternative answers is provided. The student can select one of the alternatives, following which the teacher responds in a binary manner indicating whether the selected answer is right or wrong. The teacher, however, is probabilistic—there is a nonzero probability of eliciting either of the two responses for any of the answers selected by the student. The saving feature of the situation is that it is known that the teacher’s negative responses have the least probability for the correct answer. Under these circumstances the interest is in finding the manner in which the student should plan a choice of a sequence of alternatives and process the information obtained from the teacher so that he learns the correct answer. In stochastic automaton models the stochastic automaton corresponds to the student, and the random environment in which it operates represents the probabilistic teacher. The actions (or states) of the stochastic automaton are the various alternative answers that are provided. The responses of the environment for a particular action of the stochastic automaton are the teacher’s probabilistic responses. The problem is to obtain the optimal action that corresponds to the correct answer.

The stochastic automaton attempts a solution of this problem as follows. To start with, no information as to which one is the optimal action is assumed, and equal probabilities are attached to all the actions. One action is selected at random, the response of the environment to this action is observed, and based on this response the action probabilities are changed. Now a new action is selected according to the updated action probabilities, and the procedure is repeated. A stochastic automaton acting in this manner to improve its performance is referred to as a _learning automaton_. …” (Narendra & Thathachar, 1974).

Stochasticity can take various forms (McNamara et al., 2011): “As Frank and Slatkin (1990) pointed out, stochasticity can be partitioned into variation that affects each member of a lineage independently and variance that is correlated across individuals. The former, referred to as individual variation, is also
When it acts alone, the future... (Sober 1984, p. 110;...

7. Evolutionary outcomes are probabilistic. But why?

The Oxford Dictionary defines selection as: “The action or fact of carefully choosing someone or something as being the best or most suitable”. According to this definition, selection should have a deterministic outcome. The Oxford Dictionary definition of natural selection is: “The process whereby organisms better adapted to their environment tend to survive and produce more offspring. The theory of its action was first fully expounded by Charles Darwin and is now believed to be the main process that brings about evolution.” The Darwinian concept of natural selection was conceived within a set of Newtonian background assumptions about systems dynamics. Within this conceptual framework the process of natural selection is deterministic (Sober, 1984; Brandon & Carson, 1996; Witting, 2003; Sols, 2014). This is in analogy to the breeder’s deterministic selection, Darwin’s role model of evolution. Sober (1984) elaborates: “When it acts alone, the future frequencies of traits in a population are logically implied by their starting frequencies and the fitness values of the various genotypes” (Sober, 1984, p. 110; italics in original). A role for deterministic natural selection is typically inferred when genotypes or phenotypes are similar for independent populations in similar environments: that is, parallel or convergent evolution (Endler, 1986; Schluter, 2000; Langerhans & DeWitt, 2004; Arendt & Reznick, 2008; Losos, 2011; Wake et al., 2011). As another example, specific causes of natural selection are typically inferred through correlations between genotypes or phenotypes and a particular ecological factor (Endler, 1986; Wade & Kalisz, 1990; MacColl, 2011; Hendry et al., 2013), such as diet (e.g., Schluter & McPhail, 1992;
Kaeuffer et al., 2012), structural habitat features (e.g., Losos, 2009), predation (e.g., Reznick & Bryga, 1996; Langerhans & DeWitt, 2004), or water flow (e.g., Langerhans, 2008).

There has been a general agreement that ecological and evolutionary outcomes and the tools that are used to describe them are probabilistic and statistical (Glymour, 2001; Millstein, 2003; Shanahan, 2003; Rosenberg, 2004; Colyvan, 2005). Mendelian genetics at first did not sit well with the gradualist assumptions of the Darwinian theory. Eventually, however, Mendelism and Darwinism were fused by reformulating natural selection in statistical terms. This reflected a shift to a more probabilistic set of background assumptions based upon Boltzmannian systems dynamics (Weber & Depew, 1996). This triumph was possible only in the new, more probabilistic scientific climate that, as historians of science have been arguing, began to take shape in the last half of the nineteenth century (Hacking, 1975, 1983, 1990; Krüger et al., 1987; Gigerenzer et al., 1989; Weber & Depew, 1996). It has been a contentious issue what evolutionary factors underlie the statistical character of evolutionary theory. Graves et al. (1999) argue that the probabilism of the theory of evolution is epistemically motivated. For Brandon and Carson, the probabilism of natural selection derives solely from its real-life connection with drift: “natural selection is indeterministic at the population level because in real life as opposed to certain formal models) it is inextricably connected with random drift” (1996, p. 324; italics in original). One of the most commonly encountered analogies for the process of evolution is that of a blindfolded selector drawing balls from an urn. The metaphor is thought to illuminate the irreducibly probabilistic nature of evolutionary processes (Walsh et al., 2002). Other statistical metaphors abound too: selection is spoken of as “discriminate sampling” (Beatty, 1984). Drift is spoken of variously as “indiscriminate sampling,” or “sampling error”. Accordingly, the issue of what is more important in accounting for evolutionary change, indiscriminate sampling in finite populations, or discriminate sampling is not all-or-none, but a more-or-less issue (Beatty, 1984) due to processes whose relative importance vary with population size. Natural selection, by this way of thinking, is a mere consequence of a statistical property of a population—its variation in fitness (Endler, 1986). Evolutionary theory dealt with the issue of stochasticity, both at the genetic (mutations, recombination) and population (random drift, migration) level. In the tradition of Darwin’s theory, the Modern Synthesis considered either imperfect biological processes or random drift as the sources of stochasticity. Importantly, stochasticity and natural selection are regarded as distinct entities (Millstein, 2002). Figure 2 depicts the linear evolution model as put forward in the Modern Synthesis (e.g. Mayr, 2000). Chance mutations create variation on which natural selection acts. This linear model, however, lacks a feedback loop and is unable to learn.

8. Chance and necessity

In his influential work Le hasard et la nécessité (Essai sur la philosophie naturelle de la biologie moderne) (1970), the French biologist Jacques Monod contrasts chance and natural selection as the two driving mechanisms of evolution (Sols, 2013). “Pure chance, (...) mere chance is at the very roots of the prodigious framework of evolution: today, this central biological notion (...) is the only one which is consistent with the reality shown by observations and experience” (Monod, 1971). The living world is shaped by the interplay of deterministic laws and randomness. It is widely accepted that the evolution of any particular organism or form is a product of the interplay of a great number of historical contingencies (Monod, 1971). Rewind and replay the tape of life again and again, as the now familiar argument goes, and there is no predicting (or reproducing) the outcomes (Gould, 1989). Roses and redwoods, humans and hummingbirds, trilobites and dinosaurs each owe their existence (or demise) to unfathomable combinations of innumerable rolls of the ecological and genetic dice (Carroll, 2001a). Despite the widespread occurrence and attractive mechanistic simplicity of adaptive radiations, the evolutionary outcome of an instance of adaptive radiation cannot generally be predicted with any degree of confidence. The inability to make such a prediction is due in part to an inability to evaluate the relative roles of chance and necessity (Monod, 1971) in promoting divergence (Travisano et al., 1995a, b). Longo et al. (2012) and Kaufman (2013) assert that the interactions between organisms, biological niches and ecosystems are ever changing, intrinsically indeterminate and even unprestatable. Hence, no laws of motion can be formulated for evolution (Longo et al., 2012; Kaufman, 2013). Examples of adaptive radiations suggest that either chance or adaptation can be the dominant factor in shaping the adaptive process and the resulting adaptive radiations (Wahl & Krakauer, 2000; Chan & Moore, 2002).

8.1 Chance

According to Lynch (2007a; b), out of the four major forces in evolution, natural selection, mutation, recombination and drift, the latter three are stochastic
in nature. In addition, evolutionary forces are often divided into two sorts: stochastic and deterministic (Wright, 1955).

8.1.1 Mutations
The Modern Synthesis holds that (i) mutations occur independently of the environment, (ii) mutations are due to replication errors, and (iii) mutation rates are constant (Lenski & Mittler, 1993; Brisson, 2003). Currently, biologists usually agree that all genetic mutations occur by “chance” or at “random” with respect to adaptation (Miller, 2005) and the novel allele is subsequently selected for or against. The claim dates back to Darwin’s conception of “spontaneous,” “accidental” or “chance” variation (Darwin, 1859; Darwin & Seward, 1903). The Modern Synthesis later redefined Darwin’s idea as rooted in the phenomenon of genetic mutation following a long period of controversy over the “chance” vs “directed” character of variation (Merlin, 2010). However, in the view of mathematicians, the ratio of the number of functional genes and proteins, on the one hand, to the enormous number of possible sequences corresponding to a gene or protein of a given length, on the other hand, seemed so small as to preclude the origin of genetic information by a random mutational search (see chapter 4).

Due to the high probability that any particular mutation will have deleterious effects, orthodox theory holds that “natural selection of mutation rates has only one possible direction, that of reducing the frequency of mutation to zero” (Williams, 1966). Thus, there should be a strong selective pressure to eliminate mutations altogether. Accordingly, theory indicates that under most conditions, selection puts a premium on the faithful maintenance and transmission of genetic information and is expected to favor alleles that reduce the mutation rate (Karlin & McGregor, 1974; Feldman & Liberman, 1986; Kondrashov, 1995; Sniegowski et al., 2000; Sniegowski, 2004). In fact, DNA replication can have a remarkable fidelity, estimated to produce $10^{-8}$ to $10^{-10}$ mutations/nucleotide, achieved by multiple mechanisms of error avoidance and correction (Kunkel, 2004). In well-adapted populations in stable environments the rate of mutation will evolve towards lower values (Leigh, 1973; Karlin & McGregor, 1974; Liberman & Feldman, 1986; Drake, 1991; Kunkel, 2004). Theory holds that the combined metabolic and temporal costs of perfection in replication and transcription fidelity (Kimura, 1967; Sniegowski et al., 2000) limit further improvements in replication fidelity and DNA repair (André & Godelle, 2006). Thus, stable environments would favor low mutation rates (anti-mutator genotypes), constrained only by the costs of error-repair mechanisms (Kimura, 1967; Drake, 1991). On the other hand, Eigen (1992) argued that replication error rates established themselves near an error-threshold where the best conditions for evolution exist.

There is cumulative evidence to refute the metabolite-costs-of-replication-fidelity argument. Nature is unforgiving at the edge of life (Kis-Papo et al., 2003). As a consequence of the increasingly narrower adaptive road, a variety of Archaeal extremophiles, compared to mesophiles, evolved a high genomic stability (Mackwan, 2006; White & Grogan, 2008; Kish & DiRuggiero, 2012) with low mutation rates (Battista, 1997; Grogan et al., 2001; Grogan, 2004; Mackwan, 2006; Mackwan et al., 2007; Drake, 2009), very high replication fidelity (Lundberg et al., 1991; Mattila et al., 1991; Cline et al., 1996; Grogan et al., 2001; Dietrich et al., 2002; Berkner & Lipps, 2008; Zhang et al., 2010), and decreased genetic diversity (Kis-Papo et al., 2003; Friedman et al., 2004; Sonjak et al., 2007; de los Ríos et al., 2010; Vinogradova et al., 2011). Since extreme habitats are routinely resource-limited (Waterman, 1999; 2001; Plath et al., 2007; Rampelotto, 2010; Prasad et al., 2011), the reduced mutation rates of extremophiles indicate that the perfection of replication fidelity in mesophiles is not limited by the availability of resources. Since many components of the DNA replication machinery of eukaryotes have evolved from a common ancestor in Archaea (Yutin et al., 2008), the question arises whether this loss of fidelity of the eukaryotic replication machinery has an evolutionary rationale.

8.1.2 Genetic drift
Genetic drift or allelic drift is the change in the frequency of a gene variant (allele) in a population due to chance events (Masel, 2011). Genetic drift may cause gene variants to disappear completely and thereby reduce genetic variation. Genetic drift is the stochastic fluctuation in allele frequencies caused by random differences in the fecundity and survival of individuals. Genetic drift is considered to be the most important of the stochastic forces in the evolution of natural populations (Gillespie, 2001). The term applies to many effects on populations or organisms which are said to be due to “chance” and to factors which are thought to help to produce such effects, e.g. natural disasters or “founder effects”. However, many core senses of random drift make it something which varies inversely with population size. Any strategy with non-zero reproductive fitness can persist over evolutionary time by genetic drift. As the effective population size, $N_e$ increases, genetic drift becomes weaker because the larger the population, the smaller
the proportional impact of each random event that concerns just one individual or a group of individuals. Hence, selection pressure and random drift, whose relative importance for evolution is often disputed in the literature, are equally important, although they act differently: selection promotes evolution, and random drift slows it down (Rouzine et al., 2001). In the cybernetic model of evolution, the genetic changes due to random drift are a property of the output signal that due to the iterative nature of evolution secondarily becomes an input signal.

When there are few copies of an allele, the effect of genetic drift is larger, and when there are many copies the effect is smaller. Vigorous debates occurred over the relative importance of natural selection versus neutral processes, including genetic drift. Ronald Fisher held the view that genetic drift plays at the most a minor role in evolution, and this remained the dominant view for several decades. In 1968, Motoo Kimura rekindled the debate with his neutral theory of molecular evolution, which claims that most instances where a genetic change spreads across a population (although not necessarily changes in phenotypes) are caused by genetic drift.

8.2 Necessity

Genotypic diversity enhances the evolutionary responsiveness and adaptability of populations (Ayala, 1968; Abrams & Matsuda, 1997; Yoshida et al., 2003; Gamfeldt et al., 2005; Reusch et al., 2005; Gamfeldt & Kallstrom, 2007; Becks & Agrawal, 2012; Roze, 2012) and its lack can increase extinction risk (Keller & Waller, 2002). It would be highly adaptive for organisms inhabiting variable environments to modulate mutational dynamics in ways likely to produce necessary adaptive mutations in a timely fashion. Jablonka and Lamb (2005, p.101) wrote: “it would be very strange indeed to believe that everything in the living world is the product of evolution except one thing? the process of generating new variation!” In his 1905 paper, Einstein proposed that the same random forces that cause the erratic Brownian motion of a particle also underlie the resistance to the macroscopic motion of that particle when a force is applied (Kaneko, 2009; Lehner & Kaneko, 2011). This insight can be generalized to state that the response of a variable to perturbation should be proportional to the fluctuation of that variable in the absence of an applied force (Kubo et al., 1985). In short, the more something varies, the more it will respond to perturbation, irrespective of the precise molecular details. A generalized version of the fluctuation–response relationship can be applied to evolved, dynamical systems (Sato et al., 2003; Kaneko & Furusawa, 2006; Kaneko, 2009).

The concept has been confirmed experimentally in unicellular prokaryotes and eukaryotes (Sato et al., 2003; Yomo et al., 2006; Lehner, 2010; Park et al., 2010). Metzgar and Wills (2000) argued that adaptively tuned mutation rates do not require any special foresight. Instead, they must have been selected for repeatedly in the past for their ability to generate genetic change. Mutational tuning does not require the specific generation of adaptive mutations (nonrandomness with respect to function) but rather the concentration of mutations under specific environmental conditions or in particular regions of the genome (nonrandomness with respect to time or location) (Metzgar & Wills, 2000). The literature reveals significant effects of genetic diversity on ecological processes such as primary productivity, population recovery from disturbance, interspecific competition, community structure, and fluxes of energy and nutrients. Thus, genetic diversity can have important ecological consequences at the population, community and ecosystem levels, and in some cases the effects are comparable in magnitude to the effects of species diversity (Gamfeldt et al., 2005; Fussmann et al., 2007; Gamfeldt & Kallstrom, 2007; Lankau & Strauss, 2007; Hughes et al., 2008). Moreover, theoretical and empirical studies suggest that diversity at one level may depend on the diversity at the other (Whitham et al., 2003; Abrams, 2006; Crutsinger et al., 2006; Johnson et al., 2006; Vellend, 2006; Lankau & Strauss, 2007).

9. Gedankenexperiment: evolution in stable environments

The Law of Causality states: Every event must have a cause (Hughes & Lavery, 2004). Therefore scientists explain particular events and general patterns by identifying the causal factors involved. Ordering two or more events in a causal order is crucial for a scientific understanding. Another order of events is their temporal order. While the temporal order is observable, outside of a controlled scientific experiment the causal order is not. This is because a complete causal account specifies the necessary and sufficient conditions for something to occur and both of these conditions involve counterfactual statements (Damer, 1995; Hughes & Lavery, 2004). Counterfactual statements are, by definition, not observable. But they are amenable to thought experiments. Certain disciplines such as evolutionary biology and
economics often do not lend themselves to experimentation. Although computer simulation may help to clarify issues (Casti, 1997), it remains the case that we cannot avoid frequent recourse to “thinking our way through” a problem, i.e., to thought experiment (Damper, 2006).

A living organism never enjoys a perfectly stable environment; the system to which it belongs may incur slow or quick changes that will impinge on its well-being and fitness. However, orthodox Darwinism assumes that environments are stable and traditionally, evolutionary models have assumed environmental constancy for simplicity and tractability (Keyfitz, 1977; Rubenstein, 1982; Caswell, 2001; Lee & Doughty, 2003). For example, Maynard Smith’s (1982a) often quoted book on Evolutionary Game Theory contains no reference to environmental stochasticity. The concept of environmental variance is almost completely absent from the 40 foundation papers (published from 1887–1971) identified by Real and Brown (1991). Through the 1960s, the word ‘variance’ appeared in the abstract of only about ten papers per thousand published by the Ecological Society of America (Ruel & Ayres, 1999). However, the number of such papers has increased since then to about 50 per thousand during the 1990s. This suggests a growing recognition among ecologists that an explicit consideration of variance is essential to explain many of the important patterns and processes in nature (Ruel & Ayres, 1999).

Environments display a range of instabilities. Across this range of spatiotemporal gradients of instability, adaptive responses show linear trends with regard to the generation of variation that allow to extrapolate these trends to perfectly stable environments. A general pattern emerges:

(i) In well-adapted populations in stable environments the rate of mutation will evolve towards lower values (Leigh, 1973; Karlin & McGregor, 1974; Liberman & Feldman, 1986; Drake, 1991; Kunkel, 2004).

(ii) In more stable environments, phenotypic plasticity is lost or limited because it may incur costs (Levins, 1968; Ghalambor et al., 2006; Schleicherová et al., 2013; Tonsor et al., 2013). In evolving populations of *Escherichia coli* adapting to a single nutrient in the medium, unused catabolic functions decayed and their diet breadth became narrower and more specialized (Cooper & Lenski, 2000). Adaptive plasticity is lost during long periods of environmental stasis (Masel et al., 2007). If the environment changes only very slowly relative to the generation time of the organism, then genetic specialization is favored over plasticity (Orzack, 1985); in relatively stable environments there is rather a selective pressure for the evolution of instinctive behaviors (Turney, 1996). Theory predicts diminished fitness for highly plastic lines under stabilizing selection, because their developmental instability and variance around the optimum phenotype will be greater compared to nonplastic genotypes. Theory is supported empirically: the most plastic traits exhibited heritabilities reduced by 57% on average compared to nonplastic traits (Tonsor et al., 2013). Conversely, developmental instability increases adaptive evolution in the face of changing environments (Rutherford & Lindquist, 1998; Masel, 2006). Environmental heterogeneity is the main factor for the evolution of a plastic trait (Pigliucci et al., 2006; Fusco & Minelli, 2010).

(iii) In relatively stable environments risk-spreading evolutionary strategies such as bet-hedging (see chapter 11) have a lower fitness advantage and do not pay any more (Philippi & Seger, 1989; Müller et al., 2013).

(iv) Since some cost is associated with learning (Dukas & Duan, 2000; Mery & Kawecki, 2003, 2004; Burger et al., 2008), in an absolutely fixed environment a genetically fixed pattern of behavior should evolve (“absolute fixity argument”). But if the environment is diverse and unpredictable, innate environment-specific mechanisms are of little use. Unpredictable or variable environments favor the evolution of cognition and learning (see chapter 5).

(v) Sexual reproduction is favored in intermediate stressful environments, while stable stressfree ones favor asexuality (Bürger, 1999; Moore & Jessop, 2003; Heininger, 2013) which may explain the high incidence of parthenogenesis in environments such as stable forest soils (Cianciolo & Norton, 2006; Domes et al., 2007).

In summary, stable environments select against evolvability (Altenberg, 2005) as achieved by mutagenesis, phenotypic plasticity, learning, sexual reproduction, and bet-hedging behavior. Thus, in a perfectly stable environment, evolution would virtually come to a hold, finally exploiting all beneficial mutations and maximizing individual fitness.

10. The evolutionary signature of stochastic environments

Natural environments are stochastic (Gard & Kannan, 1976; Halley, 1996; Bell & Collins, 2008; Lei, 2012). In a review of published studies on variation in recruitment, Hairston et al. (1996a) found that reproductive success of long-lived adults varied from
year to year by factors up to 333 in forest perennial plants, 4 in desert perennial plants, 591 in marine invertebrates, 706 in freshwater fish, 38 in terrestrial vertebrates, and 2200 in birds. Similarly, the recruitment success of diapausing seeds or eggs varied by factors of up to 1150 in chalk grassland annual and biennial plants, 614 in chapparal perennials, 1150 in freshwater zooplankton, and 31,600 in insects (Ellner, 1997). These figures represent the variation among years when some reproduction occurred; many of the studies also report years in which reproduction failed completely. A life-history model predicted the occurrence of skipped reproduction only for intermediate environmental qualities, with high reproductive investment being optimal at both ends of a gradient of environmental quality (Fischer, 2009). Skipped reproduction is frequently observed in nature (in fish: Bull & Shine, 1979; Engelhard & Heino, 2005; Rideout et al., 2005; Jørgensen & Fiksen, 2006; Jørgensen et al., 2006; in amphibians: Bull & Shine, 1979; Harris & Ludwig, 2004; in reptiles: Bull & Shine, 1979; Brown & Weatherhead, 2004; in birds: Ilera & Diaz, 2006). Poor individual condition and/or poor environmental quality are thought of as the main causes for skipped reproduction (Bull & Shine, 1979; Dutil, 1986; Rideout et al., 2005; Ilera & Diaz, 2006). In taxa without parental brood care, particularly insects, the number of embryos entering a habitat is usually far in excess of its carrying capacity, and larval survivorship is typically low (e.g., Berryman, 1988; Ogushi, 1991; Willis & Hendrick, 1992; Tinkle et al., 1993; Duffy, 1994; Dempster & McLean, 1998; Dixon et al., 1999) and unpredictable (Madsen & Shine, 1998; Fincke & Hadrys, 2001; Haugen, 2001; Rollinson & Brooks, 2007). Ecological factors such as deterioration of larval habitats or fluctuations in the density of food, predators, cannibals, or parasites can result in unpredictable windows of offspring survivorship (e.g., Smith, 1987; Newman, 1989; So & Dugeon, 1989; Morin et al., 1990; Messina, 1991; Anholt, 1994; Dixon et al., 1999). In insects, “while lifetime egg production is largely determined by chance” (Thompson, 1990), the numbers of mature offspring produced (fitness) is largely unpredictable (Fincke & Hadrys, 2001) and in natural populations, crucially, is poorly correlated with behavioral observations of mating, particularly for females (Thompson et al., 2011). Instead, the time span between hatching of the first and the last egg within a clutch was detected to be the most appropriate estimate of reproductive success. This was because a larger hatching span increased the likelihood that some larvae encountered a window of opportunity during which the risks of being eaten by larger conspecifics were lower (Fincke & Hadrys, 2001). Similarly, also in the water python Liasis fuscus, time of hatching, and not clutch size, was most predictive of reproductive success (Madsen & Shine, 1998). Unpredictable environmental change can lead to reduced survival, or to extinction of previously well-adapted organisms (Bell & Collins 2008; Simons, 2009). Extinction risk in natural populations depends on stochastic factors that affect individuals, and is estimated by incorporating such factors into stochastic models (Athreya & Karlin, 1971; May, 1973a, b; Gabriel & Bürger, 1992; Lande, 1993; Lynch & Lande, 1993; Ludwig, 1996; Halley & Kunin, 1999; Lande et al., 2003; Sæther et al., 2004a; Kendall & Fox, 2003; Fox et al., 2006; Melbourne & Hastig, 2008). Theory suggests that environmental stochasticity can be comparable to the accumulation of mildly deleterious mutations in causing extinction of populations smaller than a few thousand individuals (Lande, 1994, 1995, 1998).

Stochasticity can be divided into four categories, which include demographic stochasticity, the probabilistic nature of birth and death at the level of individuals (May, 1973a), environmental stochasticity, resulting in the variation in population-level birth and death rates among times or locations (Athreya & Karlin, 1971; May, 1973b), the sex of individuals (Lande et al., 2003; Sæther et al., 2004a), and demographic heterogeneity, the variation in vital rates among individuals within a population (Kendall & Fox, 2003; Fox et al., 2006). Generally, the uncertainty due to abiotic capriciousness is perceived as major source of stochasticity. Variable abiotic environments, however, are also often predictable (Beissinger & Gibbs, 1993). More than 40 years ago, Van Valen’s Red Queen hypothesis (1973a) emphasized the primacy of biotic conflict over abiotic forces in driving selection. According to the Red Queen hypothesis, each adaptation by a species is matched by counteracting adaptations in another interacting species, such that perpetual evolutionary change is required for existence. Despite continued evolution, average relative fitness remains constant: evolution is a zero-sum game (Brockhurst et al., 2014). Thus, for a vast number of biological situations, the salient aspects of the selective environment are biotic (Richardson & Burian, 1992; Venditti et al., 2010; Ezard et al., 2011; Liow et al., 2011; Brockhurst et al., 2014). For example, in human populations pathogens have a higher impact on genetic diversity than climate conditions (Fumagalli et al., 2011).

10.1 Environmental stochasticity
Spatiotemporal, often unpredictable, variation in
environmental quality is a salient feature of natural habitats (Wiens, 1976, 2000; Shorrock & Swingland, 1990; Halley, 1996; Wiens, 2000; Simons, 2002, 2009; Metcalf & Koons, 2007; Bell & Collins, 2008; Doebeli & Ispolatov, 2014). To survive to reproduce, an animal must solve multidimensional problems with components that can vary independently of one another over its lifetime. Furthermore, many of these components are fundamentally unpredictable at the spatial and temporal scales at which organisms operate (Dall, 2010). Environmental variance includes a wide array of event magnitudes – possibly a continuum spanning the spectrum from minor fluctuations occurring over short time scales (seconds or hours) to rare events leading to mass extinction. Event magnitudes are inversely related to their frequency of occurrence. Based on records of sea-level changes and temperature from the deep ocean, Steele (1985) showed that environmental variation in marine environments increases continually with longer time series over timescales from hours to millennia. The implication is that environmental events that are disproportionately influential occur at a relatively low frequency. This general pattern of variance is known as $1/f$-noise (Halley, 1996). Theoretical and empirical studies have further advanced our understanding of the structure of temporal environmental variance (Ariño & Pimm, 1995; Halley, 1996; Bengtsson et al., 1997; Cyr, 1997; Solé et al., 1997, 1999; McKinney & Frederick, 1999; Plotnick & Sepkoski, 2001).

In the Robertson–Price equation for the evolution of quantitative characters, the effects of environmental stochasticity causing fluctuating selection can be partitioned from the effects of selection due to random variation in individual fitness caused by demographic stochasticity (Engen & Saether, 2014). A stochastic version of the Price equation reveals the interplay of deterministic and stochastic processes in evolution (Rice, 2008). Demographic stochasticity can cause random variation in selection differentials independent of fluctuating selection caused by environmental variation. Populations continually evolve and interacting species continually coevolve, building a constantly coevolving web of life (Futuyma & Slatkin, 1983; Thompson, 2005, 2009) that is highly dynamic and stochastic (Dieckmann & Law, 1996; Heininger, 2013). For instance, by consuming resources, constructing nests, and excreting waste, organisms modify their environment, creating ecological feedback that alters existing selective pressures and creates others anew (Jones et al., 1994; Odling-Smee et al., 1996, 2003; Wolf et al., 1999; Laland & Sterelny, 2006; Kokko & Lopez-Sepulcre, 2007; van Dyken & Wade, 2012). Individuals from the same or different species impose selection on one another, creating a dynamically changing selective environment that evolves along with the traits that it selects (Futuyma & Slatkin, 1983; Kiester et al., 1984; Dieckmann & Law, 1996; Wolf et al., 1998). Coevolutionary pressures not only include interactions between e.g. symbionts, mutualists, pathogens and hosts, predators and prey, herbivores and plants but also density- and frequency-dependent coevolutionary interactions (Mueller et al., 1991; Doebeli & Ispolatov, 2014).

Frequency-dependent selection occurs when "the fitness of a genotype (or of an allele) is affected by its frequency within the population" (Futuyma [1986], p. 166). In some cases, a genotype is fitter when it is rare (negative frequency-dependence); in other cases, a genotype can be fitter when it is common (positive frequency-dependence) (Millstein, 2006). Frequency-dependent selection is believed to be quite common. Futuyma ([1986], p. 166) remarks that "it is likely that there is a frequency-dependent component in virtually all selection that operates in natural populations, for interactions among members of a population affect the selective advantage of almost all traits, and such interactions usually give rise to frequency-dependent effects." Frequency dependence generates an evolutionary feedback loop, because selection pressures, which cause evolutionary change, change themselves as a population’s phenotype distribution evolves, causing complicated dynamics in models (Altenberg, 1991; Nowak & Sigmund, 1993a; Gavrilets & Hastings, 1995; Schneider, 2008; Priklopil, 2012; Doebeli & Ispolatov, 2014). Density- and frequency-dependence of fitness results in a highly dynamic landscape, a fitness “sphagnum bog” (Rosenzweig, 1978; Bolnick, 2004) that is chaotic and unpredictable (Altenberg, 1991; Priklopil, 2012; Doebeli & Ispolatov, 2014) with intraspecific competition as its key driver (Milinski & Parker, 1991; Doebeli & Dieckmann, 2000).

10.2 Demographic stochasticity

While environmental stochasticity refers to situations where several individuals are affected by a common factor, demographic stochasticity refers to hazards experienced independently by each individual. It is commonly observed that even very large populations may show considerable stochastic fluctuations. In the classical birth-death population models, this is not the case. If the parameters are chosen so that the population size can be very large, these stochastic models will behave almost deterministically by the law of large numbers. The mean value of the contributions to the population change will have variance close to
Theoretical predictions proposing a link between population dynamics and individual variation (e.g. May, 1973a; Lomnicki, 1978; Leigh, 1981; Shaffer, 1981; Lande, 1993; Uchmanski, 1999) are supported by data from natural populations (Dochtermann & Gienger, 2012). Demographic heterogeneity, among-individual variation in vital parameters such as survival and reproduction, is ubiquitous (Stover et al., 2012), resulting from fine-scale spatial habitat heterogeneity (e.g., Gates & Gysel, 1978; Boulding & Van Alstyne, 1993; Menge et al., 1994; Winter et al., 2000; Franklin et al., 2000; Manolis et al., 2002; Bollinger & Gavin, 2004; Landis et al., 2005), unequal allocation of parental care (e.g., Manser & Avey, 2000; Johnstone, 2004), maternal family effect (e.g., Fox et al., 2006; Pettorelli & Durant, 2007), conditions during early development, including birth order effects (e.g., Lindström, 1999), persistent social rank (e.g., von Holst et al., 2002), and genetics (e.g., Yashin et al., 1999; Ducrocq et al., 2000; Gérdes et al., 2000; Casellas et al., 2004; Isberg et al., 2006). The stability of population sizes is related to the probability of extinction (Pimm et al., 1988; Inchausti & Halley, 2003). Individual variability in life-history traits drives demographic stochasticity and extinction risk (Dochtermann & Gienger, 2012). In some cases the presence of life-history variation doubles the persistence time of populations (Conner & White, 1999). In natural populations of waterfowl, among-individual variation contributes to more than a threefold change in survival probability (Sedinger & Chelgren, 2007). Similarly, considerable within-population variation in survival probabilities has been observed in corvids (Fox et al., 2006), lagomorphs (Rodel et al., 2004), orthopterans (Ovadia & Schmitz, 2002) and in many other systems. More generally, individual variation in life-history characteristics dramatically influences the likelihood that a population will go extinct (Kokko & Ebenhard, 1996; Kendall & Fox, 2002; Sæther et al., 2004b; Fox, 2005). Selection against demographic stochasticity, favoring reductions in variance rather than a maximization of the mean, has been invoked in the evolution of numerous life-history traits, including offspring size, offspring numbers, hatching synchrony, diapause, seed dormancy, timing of germination, timing of flowering, sex-biased dispersal, etc. (Cohen, 1966; Slatkin, 1974; Gillespie, 1977; Seger & Brockmann, 1987; Yoshimura & Clark, 1991; Lehmann & Baloux, 2007; Guillaume & Perrin, 2009; Childs et al., 2010; Simons, 2011; Gremer & Venable, 2014).


Environmental variation has long been recognized as being important in determining evolutionary patterns (Bradshaw, 1965; Levins, 1968) as well as the evolution of life histories (Murphy, 1968; Wilbur et al., 1974; Ellis et al., 2009). Much circumstantial evidence suggests that one of the main effects of natural selection has been the evolution of adaptations, such as behavioral diversification (Oster & Heinrich, 1976; Lapchin, 2002; Donaldson-Matasci et al., 2008; Starrfelt & Kokko, 2012), storage of resources (Bevison et al., 1972; Lee, 1975), increases in body size (Bell, 1971; Jarman, 1974; Boyce, 1979; Peters, 1983), and increases in mobility that buffer animals against the effects of fluctuating environments (Rubenstein, 1982). While both environmental and either phenotypic or life history variability abound, the causal relationship between both, however, may often be difficult to untangle (Lacey et al., 1983; Stearns, 1989a; Halkett et al., 2004; Viney & Reece, 2013) since changes in phenotypes or life histories may also result from processes unrelated to environmental variations or due to other nonadaptive alternatives (Stearns, 1989a; Cooch & Ricklefs, 1994; Halkett et al., 2004). Thus, it is often difficult to tell empirically whether the life history variation is produced randomly, as in bet-hedging, or in response to predictive variation...
environmental clues (Philippi, 1993b; Clauss & Venable, 2000; Adondakis & Venable, 2004; Morey & Reznick, 2004; Donaldson-Matasci et al., 2010). In some cases, it may actually be a combination of both mechanisms (Richter-Boix et al., 2006; Garcia-Roger et al., 2014).

Since the early work of Haldane and Jayakar (1963), Kimura (1965), Ewens (1967), Levins (1968) and Lewontin and Cohen (1969), it has become clear that the ability to respond to environmental variability has a selective advantage and models incorporating stochastic fluctuations in fitness are an important part of the population genetic literature (Dempster, 1955; Felsenstein, 1976; Frank & Slatkin, 1990; Jablonka et al., 1995; Lachmann & Jablonka, 1996; Ancel & Fontana, 2000). It is important to realize that, when environmental conditions fluctuate, strategies may be superior that are inferior under constant conditions (Diekmann, 2004). The evolution of learning and memory, for instance, although maladaptive in static environments is the evolutionary signature of stochastic environments (discussed in chapter 5).

Ross Ashby (1956) characterized environments and possible adaptive options using the term “variety.” His classic “Law of Requisite Variety” (1956, p. 206) holds that: “variety can destroy variety” (now commonly quoted as “only variety can absorb variety” (Beer, 1966)). His insight was that a system has to have internal variety that matches its external variety so that it can self-organize to deal with and thereby “destroy” or overcome the negative effects on adaptation of imposing environmental constraints and complexity. In biology, this is to say that a species has to have enough internal variance to successfully adapt to whatever resource and competitor tensions imposed by its environment (McKelvey, 2004a). According to a similar vein of thought (“fighting change with change” [Meyers & Bull, 2002]), a “bet-hedging strategy” (Seger & Brockmann, 1987) through the diversification of the population can cover all bases of unpredictable evolutionary scenarios. An analogy with financial problems of risk management has been noticed many times (Lewontin & Cohen, 1969; Real, 1980; Stearns, 2000; Wagner, 2003). Two distinct types of strategy exist to cope with stochastic environments: risk avoidance and risk spreading (den Boer, 1968; Seger & Brockmann, 1987; Yoshimura & Clark, 1993; Einum & Fleming, 2004). Bet-hedging involves betting so as to offset a bet already made (Diamond & Rothschild, 1978). In commerce, hedging may refer to sales of securities against previous purchases of other securities to avert possible loss or, conversely, to buy against previous sales (Boyce et al., 2002). In an uncertain market, a hedging investor can reduce the risk of devastating losses during bad times, but of course, gains during a favorable period would not be as great as if he had taken the risk. By hedging, one may reduce or eliminate risk (Boyce, 1988). Conservative strategies avoid extremes, diversified strategies offer insurance against risks (Boyce et al., 2002).

Risk avoidance, also referred to as conservative bet-hedging, is an individual adaptation. Conservative bet-hedging corresponds to pursuing a relatively slow life history strategy, in which individuals sacrifice offspring quantity for quality by producing a smaller number of offspring than would be optimal over a reproductive lifetime in a stable environment of the same average quality. The conservative strategy involves producing offspring that are reasonably well equipped to handle the range of fluctuating conditions encountered over the organism’s evolutionary history (Ellis et al., 2009). When such offspring perform fairly well across this range, and/or when environmental changes affect an entire population on the timescale of a generation (e.g., years of drought) and thus cannot be handled through niche selection, natural selection tends to favor conservative bet-hedging (Donaldson-Matasci et al., 2008).

By contrast, diversified bet-hedging is a population-level adaptation. It involves “spreading the risk” by increasing phenotypic variation among offspring, and thus increasing the probability that at least some offspring will be suited to whatever environmental conditions occur in the next generation. Diversified bet-hedging can be achieved through maintenance of genetic polymorphisms or through variable expression of phenotypes arising from a monomorphic genetic structure. When any single phenotype performs poorly across the range of changing conditions encountered over evolution (i.e., when generalist strategies fail), and/or when environments vary substantially across individuals in a single generation (enabling diverse organisms to evaluate and select niches that match their phenotypes), selection tends to favor diversified bet-hedging (Donaldson-Matasci et al., 2008).

Examples of risk spreading include genetic variation (Ellner, 1996; Sasaki & Ellner, 1997), dispersal of progeny (spatial averaging: Levin et al., 1984; Kisdi, 2002), longevity (Morris et al., 2008), iteroparity (Murphy, 1968; Bulmer, 1985; Orzack & Tuljapurkar, 1989; Wilbur & Rudolf, 2006), brood care (Bonsall & Klug, 2011, Wong et al., 2013), delayed germination of seeds (temporal averaging: Ellner, 1985), phenotypic polymorphism (Levins, 1968; Roughgarden, 1979, p. 272), canalization of genetic, developmental, and
environmental perturbations (Pfister, 1998; Proulx & Phillips, 2005), and generalism (Lachrin, 2002; Donaldson-Matasci et al., 2008; Starrfelt & Kokko, 2012). Risk spreading strategies are adaptations at the population level whereby individual members are spread or diversified into different habitats, times or strategies (Yoshimura & Clark, 1993). Organisms adapt their life histories to temporally uncertain environments with life-history delays, such as seed dormancy, variable age at maturity, iteroparity, and they adapt to spatially uncertain environments with dispersal (Wilbur & Rudolf, 2006). Several diversifying traits such as dispersal, dormancy, and seed size variation may selectively interact and are complementary and partially substitutable life-history responses to spatial and temporal environmental uncertainty (Venable & Brown, 1988). For instance, diapause and dispersal are considered as two alternative responses to unfavorable environmental conditions (Southwood, 1977; Hanski, 1988; Bohonak & Jenkins, 2003), so that temporal dispersal via developmental mechanisms (especially diapause) is considered to be functionally equivalent to spatial dispersal (Hastion, 2000; Hastion & Kearns, 2002; Bohonak & Jenkins, 2003). Diversifying bet-hedging creates variation among individuals. Structured variation among individuals in survival and fecundity can reduce demographic stochasticity (Fox & Kendall, 2002; Kendall & Fox, 2002; Fox, 2005; Fox et al., 2006; Stover et al., 2012). At its extreme, it may completely eliminate demographic stochasticity (Kendall & Fox, 2002). This feature may be a manifestation of Ashby’s (1956) ‘Law of Requisite Variety’.

Conservative and diversified bet-hedging are not mutually exclusive, and the same species may display both. Great tits (Parus major) inhabit environments characterized by substantial temporal unpredictability. One adaptation shown by them is conservative bet-hedging: Average clutch size (8.53) is below the optimal size (12), given the long-term average quality of their habitat (Boyle & Perrins, 1987). This smaller clutch size has apparently been selected for because, in bad years, individuals laying smaller clutches experience substantially better nesting success. This bad-years effect “reduces the mean and increases the variance in fitness for individuals laying large clutches more than it does for individuals laying smaller clutches” (Boyle & Perrins, 1987). Although these conditions have given rise to conservative bet-hedging, the unpredictability of the great tit’s environment has also favored diversified bet-hedging: adaptive genetic variation in personality, which can be characterized along the Hawk-Dove dimension. As reviewed by Ellis et al. (2006), unpredictable variation in climate cycles strongly affects food supplies and intrasexual competition among great tits, resulting in density-dependent selection for Hawks and Doves, but in opposite directions in good and bad years and in males and females. This covariation between the Hawk-Dove dimension of personality in great tits and fitness in fluctuating environments (Dingemans et al., 2004) provides an empirical basis for the maintenance of adaptive genetic variation as a diversified bet-hedging strategy. A single trait, e.g. flowering size, can also mediate both a conservative and diversifying bet-hedging response (Childs et al., 2010). Likewise, cooperation as bet-hedging response can be both conservative and diversifying (Fronhofer et al., 2011; Rubenstein, 2011). If the amplitude of environmental fluctuation is small enough to be covered by one phenotype, conservative bet-hedgers can evolve; otherwise diversified bet-hedgers will evolve (Yasui, 1998). It has also been argued that the traditional division between conservative and diversified strategies can be considered a false dichotomy, and is better viewed as two extreme points on a continuum (Starrfelt & Kokko, 2012). Within-generation and between-generation bet-hedging is also a false dichotomy; bet-hedging strategies can occur under any grain of the environment effectively being a combination of between-generation and within-generation characteristics (Starrfelt & Kokko, 2012).

Uncertainty can be overcome by acquiring information about an environment (Stephens, 1987, 1989); risk cannot (Winterhalder et al., 1999). To deal with uncertainty, organisms had to acquire the capacity to learn from past environments to generalize to new environments (Kirschner & Gerhart, 2005; Gerhart & Kirschner, 2007; Parter et al., 2008). Early work in population genetics (Haldane, 1957; Kimura, 1961; Felsenstein, 1971, 1978) and recent analyses of evolution in fluctuating environments (Bergstrom & Lachmann, 2004;ussell & Leibler, 2005; Donaldson?Matasci et al., 2010; Rivoire & Leibler, 2011) hint at a possible relation between information and fitness. However, evolution does not “know” in advance which evolutionary path will lead to the increase of fitness or how fluctuating, often unpredictable, environments will change (Grant & Grant, 2002). Theoretical studies show that strategies of producing random difference are best when environmental information is poor, absent or too costly to process (Perkins & Swain, 2009). Therefore, prospectively, the best strategy to increase fitness is to take every possible path at every next step. As a result, no configurations should be missed (Fu, 2007). Which configuration is a “fit” one, is finally decided by the
survival and reproductive success of the individual. Fluctuating environments can favor the evolution of mixed strategies (Haccou & Iwasa, 1995; McNamara, 1995; Sasaki & Ellner, 1995). To remain in the student-probabilistic teacher scenario (Narendra & Thathachar, 1974; see chapter 6.1), learning would progress most effectively if the student was allowed to give several alternative answers to the uncertainty at hand. Thus, the risk to miss consistently the correct answer(s), resulting in extinction of a population, would be minimized (Simons, 2007, 2008). Via the law of large numbers evolution generated a form of automatic biological insurance against idiosyncratic risk (Robson, 1996). Risk-spreading by bet-hedging can be represented by an evolutionary game (Olofsson et al., 2009).

The question whether all these variation-generating processes are accidental or are selected for amounts to the question whether bet-hedging is a haphazard process or an ESS. Evolutionary game theory (Maynard Smith & Price, 1973; Maynard Smith, 1982a) has become an important way of thinking about evolution in situations in which the fitness of particular phenotypes depends on their frequencies in the population (Parker et al., 1972; Maynard Smith, 1974; Axelrod & Hamilton, 1981; Charnov, 1982; Axelrod, 1984; Nowak & Sigmund, 1993b). The key point in Evolutionary Game Theory (EGT) models is that the success of a strategy is determined by how good the strategy is in the presence of other alternative strategies, and of the frequency that other strategies are employed within a competing population. To create a sufficient amount of winners under all realistic assumptions an evolutionary stable strategy (ESS) must ‘cover all bases’. Both theoretical and experimental approaches demonstrated that in the face of variable and unpredictable environments, bet-hedging is the ESS (Hairston & Munns, 1984; Haccou & Iwasa, 1995; Sasaki & Ellner, 1995; Beaumont et al., 2009; Olofsson et al., 2009; Rees et al., 2010; Ripa et al., 2010; Charpentier et al., 2012; Starrfelt & Kokko, 2012). In fluctuating environments it may be optimal for different individuals of the same genotype to take different actions to spread the risk and ensure the genotype is represented in future generations. It does not make sense to “put all your eggs into one basket”. What is remarkable about EGT being applicable to fluctuating environments is that the players need never physically interact, compete or even communicate; nor there be any frequency-dependent selection (Hutchinson, 1996). In lotteries, spreading of the bets is a must to improve one’s chances to win. Variation is the bet-hedging strategy to cover all bases in an often unpredictable environment. Intriguingly, theoretical modeling suggests that bet-hedging as ESS in stochastically switching systems may have a U-shaped relationship with the frequency at which the environment changes (Müller et al., 2013): (i) in systems with a rapid change, a monomorphic phenotype adapted to the mean environment, (ii) for an intermediate range, a bimorphic bet-hedging phenotype and (iii) in slowly changing environments, a monomorphic phenotype adapted to the current environment are favored. Another analysis indicated that the benefits derived from bet-hedging strategies are much enhanced for higher environmental variabilities (large external noise) and/or for small spatial dimensions (large intrinsic noise). The authors concluded that these circumstances are typically encountered by living systems, thus providing a possible justification for the ubiquitousness of bet-hedging in nature (Hidalgo et al., 2014b).

Stochasticity works as environmental stochasticity at the input level of the cybernetic Black Box and at the output level resulting from stochasticity of (i) evolutionary/molecular effector mechanisms (e.g. random drift, mutagenesis, noisy cellular gene expression) and (ii) risk-spreading response to environmental stochasticity. Both an unpredictable, fluctuating abiotic environment and constantly coevolving web of life (Grant & Grant, 2002; Thompson, 2005, 2009) contribute to stochasticity. Uncertainty can be measured as the variance of a distribution of environmental quality, and adversity as the mean (Andras et al., 2003; Fronhofer et al., 2011). Both adversity and uncertainty have been conceptualized as aspects of environmental ‘risk’ (Daly & Wilson, 2002; Dall, 2010). In response to uncertainty as to which phenotype will have highest fitness in the future, biological systems exert risk minimization by risk avoidance or risk-spreading. In general, decision theory predicts and theoretical studies show that random strategies can outperform deterministic strategies whenever some aspect of the environment is unobserved (Bertsekas, 2005; Perkins & Swain, 2009). In the face of environmental stochasticity, evolution “learned” not to “put all its eggs into one basket” but to be prepared for potential selective scenarios. The environment does not need to be variable or heterogeneous for selection to favor bet-hedging; it simply needs to create risk at all places and times (Stearns, 2000). The probability of Having Descendants Forever has been advocated as complementary to the approaches of maximizing the expected number of offspring or geometric mean growth rate (Meginniss, 1977; Levy, 2010). According to this concept, constant relative risk aversion can be...
viewed as an evolutionary-developed heuristic aimed to maximize the probability of having descendant forever.

In fluctuating environments it may be optimal for different individuals of the same genotype to take different actions to spread the risk. Risk spreading polymorphism makes sense only for groups - by definition, an individual cannot be polymorphic. The fitness of the genotype is determined by the, perhaps complementary, actions of all individuals of the genotype, and the best action of an individual depends on the states and actions of other population members (McNamara et al., 1995; McNamara, 1998; Török et al., 2004; Simons, 2009). Game-theoretic methods show that multiple strategies will coexist when types compete (Ellner, 1997). Risk minimization strategies are exerted on all levels of biological organization (Cohen, 1966; Gillespie, 1974a; Slatkin, 1974; Tonegawa, 1983; Hairston & Munns, 1984; Seger & Brockmann, 1987; Philippi & Seger, 1989; Frank & Slatkin, 1990; Moxon et al., 1994; Sasaki & Ellner, 1995; Ellner, 1997; Simovich & Hathaway, 1997; Danforth, 1999; Hopper, 1999; Menu et al., 2000; Lips, 2001; Meyers & Bull, 2002; Stumpf et al., 2002; Fox & Rauter, 2003; Friedenberg, 2003; Hopper et al., 2003; Balaban et al., 2004; Einum & Fleming, 2004; Laaksonen, 2004; King & Masel, 2007; Rollinson & Brooks, 2007; Venable, 2007; Acar et al., 2008; Gourbière & Menu, 2009; Olofsson et al., 2009; Simons, 2009, 2011; Childs et al., 2010; Monro et al., 2010; de Jong et al., 2011; Dobrzyński et al., 2011; Nicholls, 2011; Charpentier et al., 2012; Gremer et al., 2012; Morrogioniello et al., 2012; Starrfelt & Kokko, 2012; Auld & Rubio de Casas, 2013; Brutovsky & Horvath, 2013; Heiningar, 2013; Graham et al., 2014; Solopova et al., 2014). There is growing evidence of evolutionary selection for stochastic diversity-generating mechanisms in unicellular and multicellular organisms at a variety of genetic, epigenetic, developmental, and physiological levels (McAdams & Arkin, 1997; True & Lindquist, 2000; Elowitz et al., 2002; Fraser et al., 2004; Raser & O'Shea, 2004; 2005; Kaern et al., 2005; Avery, 2006; Peaston & Whitelaw, 2006; Smits et al., 2006; Lim & van Oudenaarden, 2007; Maamar et al., 2007; Acar et al., 2008; Davidson & Surette, 2008; Freed et al., 2008; Losick & Desplan, 2008; Shahrezaei & Swain, 2008; Lenormand et al., 2009; Dercole et al., 2010; Lidstrom & Konopka, 2010; Huang, 2012).

Biological fluctuations span multiple spatial and temporal scales from fast cellular and sub-cellular processes to more gradual whole-organism multi-cellular dynamics to very slow evolutionary and population-level variability. There is a continuum of bet-hedging strategies from cellular to organismal and ecological levels (Simons, 2002). There is some evidence that macroevolutionary events at or above the species level such as speciation, radiations, or extinctions (Stanley, 1998; Levinton, 2001) could be decoupled from microevolutionary ones (at the population level within species); that macroevolutionary and microevolutionary trends, at least in part, are governed by different principles (Solé et al., 1996a, 1999; Erwin, 2000; Carroll, 2001b; Plotnick & Sepkoski, 2001). The terms “microevolution” and “macroevolution” reflect the controversy (Eldredge & Gould, 1972; Stanley, 1975; Orzack, 1981; Charlesworth et al., 1982; Maynard Smith, 1989; Gould & Eldredge, 1993; Van Valen, 1994; Bennett, 1997; Erwin, 2000; Carroll, 2001; Simons, 2002) over the unity of the process of natural selection operating at different time scales. Bet-hedging theory should be considered as relevant not only to a broad range of microevolutionary studies, but may also be applied hierarchically to macroevolutionary time scales and to all phylogenetic levels. Integral to this perspective is the treatment of environmental variance as a potentially continuous variable, spanning minor fluctuations to catastrophic events, where event frequency and severity are inversely related. Recent empirical and theoretical studies suggesting the prevalence of “reddened” or “1/f” temporal spectra thus offer support to the proposed perspective (Simons, 2002).

Recent work shows stochastic switching to be a near universal feature of living systems (Kaern et al., 2005; Smits et al., 2006), arising from little other than molecular noise (Elowitz et al., 2002; Smits et al., 2006; Lim & van Oudenaarden, 2007; Maamar et al., 2007; Freed et al., 2008). In contrast, metazoan bet-hedging usually involves phenotypic diversification among an individual's offspring, such as differences in egg and seed dormancy (Hopper, 1999; Laaksonen, 2004; Evans & Denney, 2005; Evans et al., 2007; Venable, 2007; Crean & Marshall, 2009; Simons, 2009) or developmental instability (Simons & Johnston, 1997).

11.1 Molecular biological bet-hedging

Variability in biological populations is the result of many confluent factors. The most basic one is genetic diversity among individual organisms. This genetic diversity is crucial for survival of the species in an ever-changing environment (Tsimring, 2014). It has been proposed (Ferenci & Maharjan, 2014) that heterogeneity of mutational types in populations, e.g. point mutations, deletions, insertions, transpositions...
and duplications, and their flexible frequency in populations provides a source of risk avoidance and alternative evolutionary strategies. To survive in a dynamic environment, cells are equipped with gene networks that allow growth to continue in spite of changing conditions. However, this flexibility comes at a price, and cells experiencing environmental fluctuations usually do not attain their fastest growth rate. In light of this, it is likely that there are genetic mechanisms that exist because they have been selected for in natural environments, but have little competitive advantage in highly controlled laboratory experiments (Razinkov et al., 2013).

Even genetically identical organisms, such as monoclonal microbial colonies, cloned animals or identical human twins exhibit significant phenotypic variability. Traditionally, this variability was ascribed to environmental fluctuations affecting development of individual organisms (extrinsic noise), but in recent years it has become clear that significant variability persists even when genetically identical organisms are kept under nearly identical conditions (intrinsic noise) (Tsimring, 2014). Mounting experimental evidence suggests that gene expression, both in prokaryotes and eukaryotes, is an inherently stochastic process. Transcription and translation show significantly higher error rates than replication. Stochasticity can be attributed to the randomness of the transcription and translation processes (intrinsic noise), as well as to different environmental conditions or differences in the concentration of transcription factors governing the network on a cellular level (extrinsic noise) (McAdams & Arkin, 1997; Elowitz et al., 2002; Swain et al., 2002; Paulsson, 2004, 2005; Longo & Hasty, 2006; Zhuravel et al., 2010). According to mass spectrometry measurements (Ishihama et al., 2008), the median copy number of all proteins in a single E. coli bacterium is approximately 500, and 75% of all proteins have a copy number of less than 250. The copy numbers of RNAs often number in tens, and the chromosomes (and so the majority of the genes) are usually present in one or two copies. Therefore, the reactions among these species can be prone to significant stochasticity (Tsimring, 2014). Importantly, as Ashby (1956, p. 186) stated: “It must be noticed that noise is in no intrinsic way distinguishable from any other form of variety. Only when some recipient is given, who will state which of the two is important to him, is a distinction between message and noise possible.” The behavior of gene regulatory networks also displays stochastic characteristics which, in several cases, can lead to significant phenotypic variation in isogenic cell populations (Ozbudak et al., 2002; Rao et al., 2002; Cinquemani et al., 2008). For example, experimental observations suggest that stochastic uncertainty may play a crucial role in enhancing the robustness of biochemical processes (Vilar et al., 2002), or may be behind the variability observed in the behavior of biological systems (Kærn et al., 2005; Wolf et al., 2005a; Wu et al., 2005; Blake et al., 2006; Kour.et al., 2006; Cinquemani et al., 2008). Tsuda and Kawata (2010) constructed an evolutionary model of gene regulatory networks and simulated its evolution under various environmental conditions. The results showed that most features of known gene regulatory networks, particularly robustness and evolvability, evolve as a result of adaptation to unpredictable environmental fluctuations.

11.1.1 Gene expression noise

There are two main sources of uncertainty in the DNA replication process. The first has to do with which origins of replication fire in a particular cell cycle and the second with the times at which they fire (Patel et al., 2006). ‘Noise’ has been defined as an empirical measure of stochasticity (Shahrezaei & Swain, 2008). Intriguingly, at least in part, cellular noise is genetically controlled (Raser & O’Shea, 2004). Several studies suggest that gene architecture may also be an important determinant of gene expression noise (MacNeil & Walhout, 2011). The chromatin environment of a gene plays an important role in regulating stochasticity in gene expression. Histone acetylation and DNA methylation significantly affect stochasticity in gene expression, suggesting that cells are able to adjust the variability of the expression of their genes through modification of chromatin marks (Vijuelas et al., 2012). Given that the alteration of chromatin marks is itself subject to the expression of chromatin modifiers, a complex circular causality may provide the cell with many regulation loops and ultimately with a fine-tuning of its phenotype and phenotypic variability (Vijuelas et al., 2012). Genes that are controlled by promoters that possess a TATA box are noisier in their expression (Becksei & Serrano, 2000; Blake et al., 2006; Batada & Hurst, 2007; Maheshri & O’Shea, 2007; Tiros & Barkai, 2008). A strong TATA box has been shown experimentally to increase noise (Raser & O’Shea, 2005). In contrast, transcription factors known to disrupt chromatin structure correlate with low noise genes. Genes which are constitutively expressed and under an almost constant demand (commonly referred to as house-keeping genes) have below-average levels of gene expression noise. Essential proteins and proteins related to translation, the ribosome, the proteasome, and the secretory pathway exhibit low noise (Fraser et al., 2004; Bar-Even et al., 2006). Sensitivity of gene
expression to mutations increases with both increasing trans-mutational target size and the presence of a TATA box (Landry et al., 2007). Genes with greater sensitivity to mutations are also more sensitive to systematic environmental perturbations and stochastic noise. Elevated expression noise may be beneficial and subject to positive selection. Under certain conditions, expression noise increases the evolvability of gene expression by promoting the fixation of favorable expression level-altering mutations. Indeed, yeast genes with higher noise show greater between-strain and between-species divergences in expression. Elevated expression noise is advantageous, is subject to positive selection, and is a facilitator of adaptive gene expression evolution (Zhang et al., 2009). These findings provide a mechanistic basis for gene expression evolvability that can serve as a foundation for realistic models of regulatory evolution. Cell-to-cell variation in genetically identical cells of multicellular organisms is often regulated by active non-genetic mechanisms (Kimble & Hirsh, 1979; Kimble, 1981; Doe & Goodman, 1985; Sternberg & Horvitz, 1986; Priess & Thomson, 1987; Jan & Jan, 1995; Karp & Greenwald, 2003; Hoang, 2004; Colman-Lerner et al., 2005). Observations suggest that the molecular events underlying cellular physiology are subject to fluctuations and have led to the proposal of a stochastic model for gene expression and biochemistry in general (Rao et al., 2002). Other cellular processes influenced by noise include ion-channel gating (White et al., 2000), neural firing (Allen & Stevens, 1994), cytoskeleton dynamics (van Oudenaarden & Theriot, 1999) and motors (Simon et al., 1992). The generation of phenotypic heterogeneity owing to a variable gene expression depends on the genetic circuitry of a system. The specific molecular interactions and/or chemical conversions depicted as links in the conventional diagrams of cellular signal transduction and metabolic pathways are inherently probabilistic, ambiguous, and context-dependent (Kurakin, 2007). Regulatory systems or decisions, in which the outcome of a cellular event is at least partially the result of intrinsic noise, are said to be stochastic (Theise & Harris, 2006; Losick & Desplan, 2008; Eldar & Elowitz, 2010). Single-cell expression profiling experiments of high spatial and temporal resolution revealed stochastic activation of responsive genes (McAdams & Arkin, 1997; Elowitz et al., 2002; Fraser et al., 2004; Levsky et al., 2002; Raser & O’Shea, 2004, 2005).

The level of transcription of any gene is not maintained at a steady level but rather occurs as a series of rapid bursts separated by periods of lower expression (Ross et al., 1994; Newlands et al., 1998; Blake et al., 2003; Golding & Cox, 2004; Golding et al., 2005; Cai et al., 2006; Yu et al., 2006). Such bursts are entirely stochastic and occur at different times for different genes. Simulations of stochastic behavior in dynamically unstable high-dimensional biochemical networks resulted in burstiness (Rosenfeld, 2009, 2011). Using a stochastic model of simple feedback networks, Kuwahara & Soyer (2012) found that independent of the specific nature of the environment–fitness relationship, the main outcome of fluctuating selection is the evolution of bistability and stochastic switching in a gene regulatory network. Such emergence occurs as a byproduct of the evolution of evolvability and exploitation of noise by evolution (Kuwahara & Soyer, 2012). Phenotypic heterogeneity is often an outcome of gene expression dynamics involving positive feedback (Ferrell, 2002; Dubnau & Losick, 2006; Smits et al., 2006). The combined effect of positive feedback and noise provide a universal mechanism for generating phenotypic heterogeneity in cell populations (Weinberger et al., 2005; Dubnau & Losick, 2006; Kashiwagi et al., 2006; Smits et al., 2006; Karmakar & Bose, 2007; Leisner et al., 2007; Maamar et al., 2007; Suel et al., 2007; Surek et al., 2008). Positive feedback induces a switch-like behavior and bistability (Ferrell & Machleder, 1998; Ferrell, 2002; Tyson et al., 2003) and negative feedback represses noise effects (Tyson et al., 2003; Paulsson, 2004; Dublanche et al., 2006; Loewer & Lahav, 2006). Noise in gene expression does not give rise to phenotypic heterogeneity as long as it is suppressed by negative feedback but it becomes important when amplified by a positive feedback loop (Smits et al., 2006; Davidson & Surette, 2008; Sureka et al., 2008). Analysis of transcription in single cells indicated that both alleles of imprinted genes were expressed randomly, but with different probabilities (Jouvenot et al., 1999). The phenomena of monoallelic gene expression (Serizawa et al., 2003), haploinsufficiency (Cook et al., 1998) and phenotypic heterogeneity in isogenic cell populations (Blake et al., 2003) were explained by the inherently stochastic nature of gene expression (Kurakin, 2005a).

Theoretical modeling and empirical analysis of yeast data (Wang & Zhang, 2011) showed that (i) expression noise reduces the mean fitness of a cell by at least 25%, and this reduction cannot be substantially alleviated by gene overexpression; (ii) higher sensitivity of fitness to the expression fluctuations of essential genes than nonessential genes creates stronger selection against noise in essential genes, resulting in a decrease in their noise; (iii) reduction of expression noise by genome doubling offers a substantial fitness advantage to diploids over
haplolds, even in the absence of sex; (iv) expression noise generates fitness variation among isogenic cells, which lowers the efficacy of natural selection similar to the effect of population shrinkage. Thus, expression noise renders organisms both less adapted and less adaptable. Because expression noise is only one of many manifestations of the stochasticity in cellular molecular processes, the results suggest a much more fundamental role of molecular stochasticity in evolution than is currently appreciated (Wang & Zhang, 2011). If a mutation is neutral, its fixation probability is unaffected by the presence/absence of the fitness noise. The fixation probability increases with the level of fitness noise for deleterious mutations but decreases for beneficial mutations. Fitness noise also affects an allele’s time to fixation just like population shrinkage (Wang & Zhang, 2011).

Simulations show that in gene expression significant fluctuations occur on both short and long length- and timescales (van Zon et al., 2006). The fluctuations on long timescales are predominantly due to protein degradation presumably by dilution, which means that the relaxation rate of this process is on the order of 1 h (Swain et al., 2002; Paulsson, 2004; Rosenfeld et al., 2005). On much shorter length- and timescales gene expression noise is associated with the competition between repressor and RNA polymerase for binding to the promoter. When a repressor molecule dissociates from the DNA, it can rebind very rapidly: possibly on a timescale of milliseconds, or less. This timescale is much shorter than that on which the RNA polymerase binds to the promoter, which is on the order of 0.01–0.1 s. Hence, when a repressor molecule has just dissociated, the probability that a RNA polymerase will bind before the repressor molecule rebinds, is very small. A repressor molecule will on average rebind many times before it eventually diffuses away from the promoter and a RNA polymerase molecule, or another repressor molecule, can bind to the promoter. This decreases the effective dissociation rate, which increases the noise in gene expression (van Zon et al., 2006).

In addition to gene expression noise there is substantial transcription infidelity. Comparing RNA sequences from human B cells of 27 individuals to the corresponding DNA sequences from the same individuals, Li et al. (2011) uncovered more than 10,000 exonic sites where the RNA sequences did not match that of the DNA, revealing infidelity of information transmission from DNA to RNA as an additional aspect of genome variation. The number of events varied among individuals by up to sixfold across 27 subjects (Li et al., 2011). Rosenfeld et al. (2005) found that quantitative relations between transcription factor concentrations and the rate of protein production fluctuate dramatically in individual living cells, thereby limiting the accuracy with which genetic transcription circuits can transfer signals.

Related to the concept of bistability is the one of phase variation. Phase variation is a process that results in differential expression of one or more genes and results in two subpopulations within a clonal population: one lacking or having a decreased level of expression of the phase variable gene(s) and the other subpopulation expressing the gene fully (van der Woude & Bäumler, 2004; van der Woude, 2006). In specific cases, phase variation can lead to antigenic variation, for example if phase variation affects expression of a lipopolysaccharide modifying enzyme. A key feature of phase variation is that the ‘On’ and ‘Off’ phenotypes are interchangeable. Thus, a cell with gene expression in the ‘Off’ phase, that is lacking expression, retains its ability to switch to ‘On’ and vice versa. Cell switching is stochastic in the sense that no prediction can be made about which cell in a population will undergo the switch. One stochastic event can set in motion a series of events that influence the frequency of occurrence of other stochastic events. The occurrence of phase variation thus results in a heterogenic and dynamically changing phenotype of a bacterial population (Srikhanta et al., 2005; van der Woude, 2006). The term ‘contingency genes’ is often adopted to describe the class of genes that are expressed in a phase variable manner (Moxon et al., 2006). By the Oxford dictionary ‘contingency’ is defined as ‘a future event, which is possible but can not be predicted with certainty’. An encompassing view on the role of phase variation is that the generation of diverse subpopulations enhances the chance that at least one can overcome a stressful challenge, in essence a ‘bet-hedging’ strategy (van der Woude, 2006).

11.1.2 Epigenesis

DNA methylationS are not gene-locus specific and have a substantial stochastic component (Silva et al., 1993; Ushijima et al., 2003; Reiss & Mager, 2007; Raj & van Oudenaarden, 2008; Huang, 2009; Mohn & Schübeler, 2009; Feinberg & Irizarry, 2010; Petronis, 2010). The degree of fidelity in epigenetic transmission is about three orders of magnitude lower than that of DNA sequence (an error rate of 1 in 10^6 and 1 in 10^9 for DNA sequences and DNA modification, respectively) (Ushijima et al., 2003; Laird et al., 2004; Riggs & Xiong, 2004; Genereux et al., 2005; Fu et al., 2010; Petronis, 2010). The cardinal signs of epigenetic effects on gene transcription are variable expression...
of a gene in a population of isogenic individuals (variable expressivity) and/or a mosaic pattern among cells of the same type within an individual (variegation) (Whitelaw & Martin, 2001). Stochastic epigenetic variation plays an important role for adaptation to fluctuating environments: by modifying the geometric mean fitness (see chapter 15.3), variance-modifying genes can change the course of evolution and determine the long-term trajectory of the evolving system (Carja et al., 2013).

11.1.3 Protein promiscuity

There is a growing realization that proteins are not as ‘ligand specific’ as the textbooks, or crystal structures, suggested (Atkins, 2014). It often happens that a polypeptide chain’s free energy conformational space does not have a well-defined minimum; this may result in markedly different structures and chemical behaviors of proteins after folding (Grosberg, 2004). Functional promiscuity (a great number of proteins can interact with a great number of molecular partners) or ‘messiness’ of most enzymes, receptors or other proteins has clear roles in biology (Glasner et al., 2007; Noy, 2007; Basu et al., 2009; Nobeli et al., 2009; Tokuriki & Tawfik, 2009; Khersonsky & Tawfik, 2010; Tawfik, 2010; Giuseppe et al., 2012; Mohamed & Hollfelder, 2013). It has been argued that, because proteins lack specificity, biological molecular interactions are, by themselves, intrinsically stochastic (Bork et al., 2004; Kupiec, 2009; Atkins, 2014). They are subject to large combinatorial possibilities making simple auto-assembly insufficient for the explanation of ontogenesis. This stochastic phenomenon is different from noise. It is not due to solely fluctuations in the concentration of molecules present in small number but to the lack of specificity of proteins causing widespread competition between them for interaction. Taking into account the intrinsically stochastic behavior of proteins necessarily brings genetic determinism into question (Kupiec, 2010). Promiscuous intermediates are highly evolvable and it has been suggested that promiscuity is actually selected as an advantageous trait within the entire proteome in order to ensure evolutionary adaptability. In fact, significant experimental evidence suggests that protein/enzyme promiscuity per se is a trait that is required to optimize evolutionary efficiency, because fewer mutations may be required when starting from a promiscuous template than from a previously optimized enzyme with high specificity (Williams et al., 2007; Chakraborthy, 2012; Tokuriki et al., 2012; Dellus-Gur et al., 2013; Díaz Arenas & Cooper, 2013; Atkins, 2014).

11.1.4 Energy-Ca^2+-redox triangle

The ultimate carriers of molecular biological stochasticity are the agents of the energy-Ca^2+-redox triangle that is modulated e.g. by metabolic stress due to maladaptation (Brookes et al., 2004; Camello-Almaraz et al., 2006; Feissner et al., 2009; Peng & Jou, 2010). Evidence is accumulating that the environment is able to shape the phenotype of organisms not only by the action of intragenerational natural selection but also by transgenerational processes (Jablonka & Lamb, 1995, 2005, 2007; Caporale, 1999, 2003a, b, 2009; Radman et al., 1999; Shapiro, 2011; Heininger, 2013). In a random environment, transgenerational effects deliver higher fitness than either a plastic only or genetic only strategy (Jablonka et al., 1995; Hoyle & Ezard, 2012). The adaptive stress in a given environment determines the metabolic condition of organisms that establishes a feedback loop for the fit between environmental and (epi)genotypic/phenotypic condition (see Heininger, 2013). This flow of information is not coded and specific as from gene to protein but code-free and stochastic. The randomness of the feedback from environment to the genome relies on the simple, codeless messenger agents ATP, Ca^2+ and free radicals (Saran et al., 1998), both regulated by and regulating cellular and organismal homeostasis in a feedback triangle (Brookes et al., 2004; Camello-Almaraz et al., 2006; Yan et al., 2006; Feissner et al., 2009; Kowaltowski et al., 2009). Cellular oxidative stress-dependent responses, although undoubtedly programmed, are also highly variable (Heininger, 2012, 2013), at least in part based on the stochasticity of mitochondrial bioenergetic/oxidative events (Hüser et al., 1998; Genova et al., 2003; Passos et al., 2007; Wang et al., 2008). In addition to cellular processes, these agents regulate organismal life history events like development and aging (Heininger, 2012) in response to environmental cues. The regulated stochastic nature of the effectors and the degeneracy of (epi)mutagenic tools (Edelman & Gally, 2001; Whitacre & Bender, 2010) may act both as a source of robustness and evolvability (Heininger, 2013). These stochastic factors allow multiple solutions for a given problem (Lenski & Travisano, 1994; Rosenzweig et al., 1994; Finkel & Kolter, 1999) and therefore have given rise to the huge diversity of evolution with an ever increasing complexity (Adami et al., 2000).

11.2 Phenotypic and behavioral bet-hedging

Bet-hedging is the risk-minimizing response to environmental uncertainty both at the individual and population level (Simovich & Hathaway, 1997; Einum & Fleming, 2004; Marshall et al., 2008; Beaumont et
Phenotypic plasticity, that is, the ability of a genotype to develop different phenotypes in different environments (Stearns 1989a), is an important characteristic that is subject to natural selection (Halkett et al., 2004). Phenotypic plasticity in insects has usually been equated with predictive plasticity, or conditional polyphenism (Walker, 1986), in which a genotype responds to different current environments by producing different phenotypes in a way that maximizes its fitness. The term “conditional polyphenism” describes this form of plasticity corresponding, for a single genotype, to a response by a single genotype to a given current environment by deterministically producing a given phenotype. However, the “decision” made now will generally have consequences on future fitness although the future state of the environment cannot be perfectly predicted on the basis of the current one. Thus, there may be a delay between the instant when the decision is made and the instant when it affects individual fitness, and during this delay the environment may change (Moran, 1992). In such a case, a stochastic decision, called adaptive coin-flipping (Cooper & Kaplan, 1982; Kaplan & Cooper, 1984; Halkett et al., 2004) or stochastic polyphenism (Walker, 1986), can be litter (Cooper & Kaplan, 1982; Haccou and Iwasa 1995; Menu et al., 2000) and can lead to diversified bet hedging (Seger & Brockman, 1987; Hopper, 1999; Menu et al., 2000; Menu & Desouhant, 2002). The term “stochastic polyphenism” describes the form of plasticity corresponding to a response by a single genotype to a given current environment by stochastically (e.g., flipping a coin) producing one phenotype among a set of possible phenotypes (Plantegenest et al., 2004). Because an organism can never perceive its environment with complete accuracy, all decision making is made under some uncertainty, and this frequently leads to a selective advantage for genotypes performing stochastic polyphenism. This form of plasticity, which copes with uncertainty, can thus be expected to be widespread in nature (Walker, 1986; Moran, 1992; Halkett et al., 2004). In models, the phenotypes that are not variable are outcompeted by those able to generate variation and innovations (Fontana & Schuster, 1998; Ancel & Fontana, 2000; Meyers et al., 2005). Fluctuations in the physical environment may even be drivers of evolutionary transitions (Boyle & Lenton, 2006).

The general question of whether one should expect environmental fluctuations to select for a component of randomness in the expression of phenotypic plasticity has received rather little attention (but see Walker, 1986; Haccou and Iwasa, 1995; Van Dooren, 2001; Koops et al., 2003; Halkett et al., 2004; Crean & Marshall, 2009; Simons, 2009; Charpentier et al., 2012). Phenotypic diversification has been hypothesized to be a form of bet-hedging, a survival strategy analogous to stock market portfolio management. From this point of view, “selfish” genotypes diversify assets among multiple stocks (phenotypes) to minimize the long-term risk of extinction and maximize the long-term expected growth rate in the presence of (environmental) uncertainty (Wolf et al., 2005a). Soil and sediment banks of dormant propagules result in overlapping generations and a prolonged generation time for an otherwise short-lived organism. It may also result in the reintroduction of genotypes which may have done poorly in previous years and a constant reshuffling of genotypes with variable past success (Templeton & Levin, 1979; Hairston & DeStasio, 1988; Ellner & Hairston, 1994; Simovich & Hathaway, 1997; Evans & Dennehy, 2005; Evans et al., 2007). Typically in unfavorable environments, some organisms have environmentally induced arrested development at different stages: embryonic diapause (Moriyama & Numata, 2008), larval diapause (Golden & Riddle, 1984), and pupal diapause (Belozerov et al., 2002). Stochastic parsing of viral populations into lytic and lysogenic (or latent) states is believed to have evolved as an adaptive solution to fluctuations in the
availability of bacterial hosts (Mittler, 1996; Stumpf et al., 2002). In the phage lambda infection process, which is governed by the lysis–lysogeny decision circuit, only a fraction of infecting phage chooses to lyse the cell. The remainder become dormant lysogens awaiting bacterial stress signals to enter the production phase of their life cycle (Ptashne, 1998).

Dispersal phenotypes could be subject to bet-hedging as well; when an environment consists of niches that become available stochastically for colonization, the optimal genotype produces a mix of dispersing and non-dispersing progeny (Comins et al., 1980). Bet-hedging in the plant kingdom might also be common, as exemplified by the probabilistic germination strategies favored by desert plants subjected to random rain-drought patterns (Satake et al., 2001). In a large range of taxa, including plants, insects, fishes and birds, offspring size variability has been suggested to be of adaptive value in variable habitats (McGinley et al., 1987; Geritz, 1995; Geritz et al., 1999; Moles & Westoby, 2006; Westoby et al., 1996). As egg phenotype is linked to offspring phenotype, increased within-brood variation in egg phenotype can have a selective advantage in unpredictable environments by increasing maternal geometric fitness (Marshall et al., 2008; Crean & Marshall, 2009; Crean et al., 2012). There is less egg size variability, both within and among female brook trouts, when environments are more predictable, and females use variability in egg size to offset the cost of imperfect information when producing smaller eggs (Koops et al., 2003). Even in clonal plants the production of variably sized offspring has been shown to be adaptive to temporal variability in environmental conditions (Charpentier et al., 2012).

Almost all known microbial bet-hedging strategies rely on low-probability stochastic switching of a heritable phenotype by individual cells in a clonal group (Thattai & van Oudenaarden, 2004; van der Woude & Bäumler, 2004; Kussell et al., 2005; Russell & Leibler, 2005; Wolf et al., 2005a; Avery, 2006; Moxon et al., 2006; Smits et al., 2006; Veening et al., 2008a, b; Beaumont et al., 2009; Fraser & Kaern, 2009; Gordon et al., 2009; Ratcliff & Denison, 2010; de Jong et al., 2011; Libby & Rainey, 2011; Rainey et al., 2011; Levy et al., 2012). The capacity to switch stochastically between heritable phenotypic states is common in bacteria (Libby & Rainey, 2011). Observed initially as variation in the morphology of colonies arising from single clones of certain bacterial pathogens (Andrewes, 1922), adaptive stochastic phenotype switching has been identified e.g. in (i) the case of bacterial persistence. Cells switch stochastically between growing and non-growing (persister) states (Keren et al., 2004) that can be adaptive in the face of periodic encounters with antibiotics despite the cost associated with non-growing cells (Balaban et al., 2004; Kussell & Leibler, 2005; Gefen & Balaban, 2009); (ii) the competence to non-competence switch for natural DNA transformation in the soil bacterium Bacillus subtilis (Maamar et al., 2007). Like the persister state, competence is associated with periods of nongrowth in an otherwise growing population and can be beneficial, despite the cost, provided the population periodically encounters conditions that kill growing cells (Johnsen et al., 2009); (iii) Haemophilus influenzae avoidance of recognition by the host immune response. H. influenzae experiences unpredictable environmental fluctuations in terms of host immune response with varying dynamics and degrees of uncertainty; whether or not bet hedging evolves depends on many factors (Thattai & van Oudenaarden, 2004; Kussell & Leibler, 2005; Kussell et al., 2005; Wolf et al., 2005b; King & Masel, 2007; Acar et al., 2008; Donaldson-Mata saci et al., 2010; Gaal et al., 2010), including the existence and reliability of environmental cues (Bull, 1987; Donaldson-Mata saci et al., 2008), the capacity of the population to respond via mutation and selection (King & Masel, 2007; Visco et al., 2010), the nature of the fitness landscape (Salathé et al., 2009; Gaal et al., 2010) and the cost–benefit balance of different strategies (Kussell & Leibler, 2005; Kussell et al., 2005; Gaal et al., 2010; Visco et al., 2010). The de novo evolution of a bet-hedging or risk-reducing strategy evolved in bacteria by a selective regime that captured essential features of the host immune response. The experimental regime involved strong frequency-dependent selection realized via dual imposition of an exclusion rule and population bottleneck (Beaumont et al., 2009; Libby & Rainey, 2011; Rainey et al., 2011).

11.2.1 Canalization and phenotypic plasticity: two sides of the same coin

The genotype-phenotype map is the common theme underlying such varied biological phenomena as genetic canalization, developmental constraints, biological versatility, developmental dissociability, and morphological integration (Wagner & Altenberg, 1996). Environmental canalization and phenotypic plasticity represent features of either conservative or diversifying bet-hedging. Environmental canalization is the insensitivity of a phenotype to variation in the environment; in the broad sense, environmental canalization refers to any kind of robustness against environmental perturbations (Waddington 1942, 1957; Roff, 1997; de Visser et al., 2003; Platt, 2005). For
example, the stimulation of a stress response can reduce mutation penetrance in Caenorhabditis elegans. Moreover, this induced mutation buffering varies across isogenic individuals because of interindividual differences in stress signaling (Casanueva et al., 2012). In contrast, phenotypic plasticity is the sensitivity of the phenotype produced by a single genotype to variation in the environment (e.g., Bradshaw, 1965; Stearns 1989b; Roff, 1997). Thus, environmental canalization and phenotypic plasticity describe different aspects of the same phenomenon: the dependency of the phenotype on the environment (Roff, 1997; Ancel & Fontana, 2000; Rutherford, 2000; Debat & David, 2001; de Visser et al., 2003; Proulx & Phillips, 2004). Importantly, canalization and plasticity are not mutually exclusive (Stearns & Kawecki, 1994). With regard to reaction norms, canalization is characterized by flatter, and plasticity by steeper, slopes of environmental sensitivity (Falconer, 1990; de Jong, 1990). Theoretical and empirical studies showed that the level of phenotypic plasticity/canalization in a trait and variation in the slope of reaction norm are under selection and dependent on levels of temporal and spatial environmental heterogeneity (Scheiner, 1993; Ellers & van Alphen, 1997; Pfennig & Murphy, 2002; Price et al., 2003; Hassall et al., 2005; Pigliucci et al., 2006; Winterhalter & Mousseau, 2007; Liefting et al., 2009; Fusco & Minelli, 2010). If environmental change is recurring, predictable, or sufficiently gradual, then adaptive phenotypic plasticity is expected to evolve (Via & Lande, 1985; Thompson, 1991; Leimar et al., 2006; Reed et al., 2010; Beldade et al., 2011; Graham et al., 2014). In a computer simulation model, environmental variation and uncertainty affect whether or not plasticity is favored with different sources of variation—arising from the amount and timing of dispersal, from temporal variation, and even from the genetic architecture underlying the phenotype—having contrasting, interacting, and at times unexpected effects (Scheiner & Holt, 2012).

Fluctuating environments generate traps on the fitness landscape. This phenomenon has been modelled in the context of drug resistance and compensatory mutation (Tanaka & Valckenborgh, 2011). The authors suggested that this phenomenon may be found more widely in nature than in the context of drug resistance and compensatory mutation. For example, cryptic genetic variation can accumulate in a genome, whose selective effects are unveiled when the environment changes (Gibson & Dworkin, 2004; Rouzic & Carlborg, 2008). The dynamics of crossing a fitness valley can be regarded as a stochastic process; the rate of traversal is a function of various parameters underlying the population biology (Stephan, 1996; Carter & Wagner, 2002; Weinreich & Chao, 2005; Durrett & Schmidt, 2008; Gokhale et al., 2009; Weissman et al., 2009; Lynch & Abegg, 2010). In addition to the steepness of the valley—the selective coefficients—the population size has been identified as an important parameter.

11.2.1.1 The Hsp90 protein at the hub of the canalization-plasticity balance

The Hsp90 stress response protein is the major molecular biological hub of the canalization-plasticity balance (Taipale et al., 2010; Heininger, 2013). Hsp90 is an ancient, abundant and nearly ubiquitous protein chaperone that interacts in an ATP-dependent system with more than 100 ‘client proteins’ in the cell, most of which are involved in signaling pathways, including protein kinases, transcription factors and others, and either facilitates their stabilization and activation or directs them for proteasomal degradation. Hsp90 is extremely abundant—constituting ~1% of total protein under normal growth conditions—and these levels may even increase following environmental stress up to tenfold both in prokaryotes and in eukaryotes (Buchner, 1999). Complete loss of Hsp90 function is lethal, as multiple essential pathways are inactivated. By linking genetic variation to phenotypic variation, the Hsp90 protein folding reservoir might promote both stasis and change (Jarosz & Lindquist, 2010). The Hsp90 chaperone system alters relationships between genotypes and phenotypes under conditions of environmental stress (Rutherford & Lindquist, 1998; Sangster et al., 2008; Jarosz & Lindquist, 2010; Chen et al., 2012) and, in so doing, provides at least two routes to the rapid evolution of new traits: (i) Acting as a potentiator. Hsp90’s folding reservoir allows individual genetic variants to immediately create new phenotypes; when the reservoir is compromised, the traits previously created by potentiated variants disappear. (ii) Acting as a capacitor, Hsp90’s excess chaperone capacity buffers the effects of other variants, storing them in a phenotypically silent form; when the Hsp90 reservoir is compromised, the effects of these variants are released, allowing them to create new traits. The loss of Hsp90 function under high stress may be due to its ATP-dependent functioning when ATP becomes limiting energetic stress-dependently (Panaretou et al., 1998; Buchner, 1999; Rutherford et al., 2007). Moreover, ROS-dependent degradation of Hsp90 protein may result in the loss of Hsp90 chaperone function, leading to client protein degradation, possibly by an ADP- and iron-dependent local generation of hydroxyl radicals through a Fenton-type reaction (Beck et al., 2012).
Thus, canalization and phenotypic plasticity are two sides of the same coin. Under environmental stress the function of Hsp90 breaks down affecting the odds for a change of the redox-dependent canalization-phenotypic plasticity balance. Hsp90 can be considered one of the key regulators of evolvability (Wagner et al., 1999; Milton et al., 2006; Rutherford et al., 2007).

11.3 Stress and bet-hedging

Adversity has the effect of eliciting talents, which in prosperous circumstances would have lain dormant. Horace (65BC-6BC)

While the causal relationship between environmental stress and bet-hedging behavior may be hard to establish at the population level, the stress-bet-hedging relationship is amenable to experimental manipulation at the molecular biological level. And here the picture is nonambiguous: stress causes molecular bet-hedging. Stress is here defined as an environmental condition to which organisms are poorly adapted and that reduces Darwinian fitness (Sibly & Calow, 1989; Zhivotovsky, 1997; Rion & Kawecki, 2005). Environmental stress is one of the most important sources of natural selection, as is witnessed by many specific adaptations evolved to alleviate the consequences of stress (e.g. Hoffman & Parsons, 1991; Randall et al., 1997; Heininger, 2001). Importantly, what is perceived as stressor depends on the evolutionary and ecological history of an organism, a change in the usual environmental conditions for any given life form. A certain environment may be claimed as stressful only if considered with respect to both a given population and the environment in which the population has evolved (Zhivotovsky, 1997; Bijlsma & Loeschcke, 2005). It follows that while a specific condition (e.g. a temperature of 65 °C) may be stressful (or even lethal) to a certain microorganism that normally lives at 37 °C, it will be optimal for growth to a thermophilic organism (Conway de Macario & Macario, 2000). As an extreme example, the deep-sea barophilic hyperthermophile Thermococcus barophilus obviously experiences stress when grown under atmospheric pressure (Marteinsson et al., 1999).

Similarly, abundant resources appear to be stressful for human populations that have been selected for their thrifty genotype (Neel, 1962; Editorial, 1989; Hales & Barker, 1992; Allen & Cheer, 1996; Fernández-Sánchez et al., 2011). Intrapopulation diversity is a mechanism to ensure survival upon exposure to environmental challenges (Booth, 2002; Avery, 2006). Bacillus subtilis responds to environmental stress with an arsenal of probabilistically invoked survival strategies (Msadek, 1999; Maughan & Nicholson, 2004). Only a certain fraction of cells within a population actually embark on a particular pathway. These include synthesis of extracellular polymer-degrading enzymes (Msadek et al., 1993), competence for DNA uptake (Hadden & Nester, 1968; Solomon & Grossman, 1996; Dubnau & Lovett, 2002), motility and chemotaxis (Frederick & Helmann, 1996; Mirel et al., 2000), biofilm and fruiting body formation (Branda et al., 2001), adaptive mutagenesis (Sung & Yasbin, 2002), and cellular differentiation into dormant and resistant spores (Perego & Hoch, 2002; Piggot & Losick, 2002). Random fluctuations in the concentration of certain key regulator molecules (Chung et al., 1994; Grossman, 1995; Dubnau & Lovett, 2002; Perego & Hoch, 2002) among individual cells may result in phenotypic fragmentation of populations. The proximate decision to embark on the different developmental pathways is not encoded in the genome, as none of five populations of B. subtilis showed any response to selection after over 5,000 generations of directional selection for enhanced efficiency of spore formation (Maughan & Nicholson, 2004). Stochastic differentiation into a growth-arrested but stress-resistant state (such as a spore) may optimize survival in an uncertain, frequently stressful environment by segregating two essential tasks: growth in the absence of stress and survival in the presence of stress (Balázsi et al., 2011). Theory has shown that a population of cells capable of random phenotypic switching can have an advantage in a fluctuating environment (Thattai & van Oudenaarden, 2004; Kussell & Leibler, 2005; Wolf et al., 2005a). Recent experiments confirmed these predictions, showing that noise stabilizes molecular network architecture under stress (Bollenbach & Kishony, 2009; Ça?atay et al., 2009), can aid survival in severe stress (Booth, 2002; Blake et al., 2006; Bishop et al., 2007), and optimize survival in specific fluctuating environments (Acar et al., 2008).

Recent findings in yeast suggest an intricate relationship between growth rate, stress resistance and bet-hedging strategies (Levy et al., 2012). As is true in descriptions of bacterial bet-hedging and persistence (Balaban et al., 2004), slow growth is a crucial predictor of stress survival in yeast (Levy et al., 2012). Both bacteria and yeast appear to be maximizing population fitness by balancing fast growth in good conditions with bet-hedging against bad ones (Kussell & Leibler, 2005). Both trehalose content and expression levels of Tsl1, a trehalose-synthesis regulator, are correlated with resistance to various forms of stress, including heat, freezing, desiccation, and high ethanol content (Hottiger et al., 1987; Crowe
et al., 1992; Winderickx et al., 1996; Singer & Lindquist, 1998; Kandror et al., 2004; Bandara et al., 2009), and growth rate in yeast (Levy et al., 2012). In contrast to bacteria, where persisters and non-persisters constitute binary growth states that predict survival in an all-or-none fashion (Balaban et al., 2004), populations of yeast might contain a continuum of metastable epigenetic cell states that each confer a different fitness in a given environment (Levy et al., 2012). Moreover, while the vast majority of characterized bacterial two-state systems are thought to interconvert through a stochastic mechanism (Hernday et al., 2003, 2004; Balaban et al., 2004; Fujita & Losick, 2005; Maamar & Dubnau, 2005; Smits et al., 2005; Veening et al., 2005; Maamar et al., 2007), differences in yeast growth and survival appear to be due to a more complex combination of stochastic and deterministic factors (Levy et al., 2012).

Further evidence for the view that microbes are single-celled stockbrokers (Wolf et al., 2005a) comes from observations that stress phenotypes introduce a trade-off between a fitness advantage under stress with a fitness defect under more favorable conditions (Cooper & Lenski, 2000; Kishony & Leibler, 2003). Diversification could be a response to this trade-off ensuring the availability of ‘favored’ phenotypes for growth in each environmental condition. Bet-hedging may involve the production of fewer and larger offspring (conservative) or of variable-sized offspring (diversified). Conservative bet-hedging strategies are recognized by a reduction in individual-level variance in fitness while diversified bet-hedging strategies are recognized by a reduction in between-individual correlations in fitness (Starrfelt & Kokko, 2012). In relatively stable environments bet-hedging strategies have a lower fitness advantage and do not pay any more (Philipp & Seger, 1989; Müller et al., 2013).

Importantly, variation is created condition-dependently, when variation is most needed— in organisms under stress. A variety of cellular noisy processes are rendered even noisier under conditions of stress. Thus, stress elicits increased mutagenesis, increased epimutagenesis, increased recombination, increased transposon mobility, increased repeat instability, increased phenotypic plasticity, and, in organisms that can reproduce both asexually and sexually, increased sexual reproduction (Heininger, 2013).

1. Biological systems function robustly despite uncertainty due to stochastic phenomena, fluctuating environments, and genetic variation (Mcd Adams & Arkin, 1999; Stelling et al., 2004). The regulation and expression of some genes are highly robust; their expression is controlled by invariant expression programs. For example, in development and differentiation, little deviation is tolerated. However, responses to stress can be more stochastic. Genome-scale studies in yeast have shown that while dose-sensitive genes and proteins forming multicomponent complexes tend to have low gene expression noise (Fraser et al., 2004; Batada & Hurst, 2007; Lehner, 2008), stress-related genes and proteins responding to changes in the environment tend to display high noise (Bar-Even et al., 2006, Newman et al., 2006; Fraser & Kaern, 2009; Stewart-Orrinstein et al., 2012). Generally, gene expression noise increases following cellular stress and oxidative stress (Thattai & van Oudenaarden, 2004; Bahar et al., 2006; Nei dez-Nguyen et al., 2008) that may lead to random cell fates at random times (Chang et al., 2008; Raj & van Oudenaarden, 2008; Gandrillon et al., 2012). Gene expression noise can be exploited to generate a fitness advantage under stress (Booth, 2002; Avery, 2006; Smits et al. 2006; Lopez-Maury et al., 2008; Losick & Desplan, 2008; Fraser & Kaern 2009; Zhuravel et al., 2010). Severe stress causes a global increase in gene expression noise in Escherichia coli (Guido et al., 2007), and increased extrinsic noise in Bacillus subtilis is used to trigger phenotypic switching in response to stress (Maamar et al., 2007). Mycobacterial survival under stress via the stringent response is dependent on positive feedback- and noise-driven bistability transitions (Sureka et al., 2008, see chapter 11.1.1). Stern et al. (2007) reported that yeast cells adapt to novel challenges by undergoing global transcriptional reprogramming that involves random gene activation, as the changes in expression of most genes were irreproducible in repeat experiments. The authors proposed a general adaptive strategy that would allow cells to overcome a broad range of stress environments by mediating stochastic gene activation (Stern et al., 2007). By broadening the range of environmental stress resistance across a population, added gene expression noise could increase the likelihood that some cells within the population are better able to endure environmental assaults (Booth, 2002; Avery, 2006). Experimental results providing support for this hypothesis were obtained in a study by Bishop et al. (2007) who demonstrated a competitive advantage of stress-resistant yeast mutants under high stress due to increased phenotypic heterogeneity. The most prominent biological examples demonstrating the benefits of stochasticity in phenotypic diversification, including persistence, sporulation and competence, represent bet-hedging strategies. In these systems, stochasticity increases phenotypic diversity in anticipation of a future adversity at the expense of reduced mean fitness (Fraser & Kaern, 2009). These observations suggest two possible stress-response mechanisms where high extrinsic noise plays a constructive role; one where it generates phenotypic diversity by increasing the variability in downstream gene expression, and one where it serves as a stochastic trigger of stress response programs (Zhuravel et al., 2010).

2. Species-wide depletion of accessible beneficial mutations requires a degree of environmental constancy that is not typical of the earth's history (Lambeck & Chappell, 2001; Zachos et al., 2001;
Eldredge et al., 2005). Genetic diversity is crucial for survival of a species in an ever-changing environment. Thus, error-free DNA repair may be maladaptive in mutagenic or stressful environments (Brevik & Gauderman, 2004; Ponder et al., 2005; Siegel-Cachedenier et al., 2007; Zhao & Epstein, 2008; Heininger, 2013). Mathematical models suggest that mutation rates adapt up (or down) as the environmental demands for novelty in variable environments (genetic innovation) or memory in stable environments (genetic conservation) increase (Bedau & Packard, 2003; Buchanan et al., 2004; Clune et al., 2008; Dees & Bahar, 2010). The optimal genomic mutation rate was found to depend only on the environmental change and its severity (Nilsson & Snoad, 2002; Ancliff & Park, 2009). A generic thermodynamical analysis of genetic information storage yielded the insight that mutation rate depends on availability/utilization of metabolic resources. A lowered ability to employ metabolic resources in mutation suppression increases the minimum effective mutation rate. This predicts transient mutation rate increase as a response to stress (Hilbert, 2011). Even in stable abiotic environments, relatively high mutation rates may be observed for traits subject to cyclical frequency-dependent population dynamics (Allen & Scholes Rosenbloom, 2012). Stress-induced mutation is a collection of molecular mechanisms in bacterial, yeast and animal cells that promote mutagenesis specifically when cells are maladapted to their environment, i.e. when they are stressed. In this sense, stress-induced bacterial and eukaryotic mutagenesis (Sung & Yasbin, 2002; Loewe et al., 2003; Achilli et al., 2004; Tenaillon et al., 2004; Galhardo et al., 2007; Robleto et al., 2007; Pybus et al., 2010; Rosenberg, 2011; Shee et al., 2011a, b; Heininger, 2013) are bet-hedging behaviors.

3. A major part of epigenetic variation is triggered by (oxidative) stress or changes in the environment (Breitranangkoon et al., 1997; Finnegan, 2004; Labra et al., 2002; Wada et al., 2004; Rapp & Wendel, 2005; Grant-Downton & Dickinson, 2006; Richards, 2006; Choi & Sano, 2007; Bossdorf et al., 2008; Boyko & Kovalchuk, 2008, 2011; Mason et al., 2008; Jablonka & Raz, 2009; Turner, 2009; Angers et al., 2010; Halfmann & Lindquist, 2010; Verhoeven et al., 2010; Flatscher et al., 2012; Grätzlov et al., 2012; Heininger, 2013). Epigenetics is closely linked to cellular bioenergetics (Xie et al., 2007; Naviaux, 2008; Smiraglia et al., 2008; Minoccheromji et al., 2012) and is, at least in part, regulated by mtDNA copy number and mitochondrial energetics (Heininger, 2013). Adverse environmental conditions play a key role in transgenerational inheritance. Conditions of stress seem to be particularly important as inducers of heritable epigenetic variation, and lead to changes in epigenetic and genetic organization that are targeted to germline specific genomic sequences (Heininger, 2013). Moreover, if conditions return to their original state, spontaneous back-mutation of epialleles can restore original phenotypes [e.g., in position-effect variegation (Richards, 2006; Flatscher et al., 2012)].

4. Increased recombination has been observed in response to stress (fitness-associated recombination) (Plough, 1917, 1921; Grell, 1971, 1978; Zhuchenko et al., 1986; Parsons, 1988; Gessler & Xu, 2000; Hadany & Beker, 2003a, b; Schoustra et al., 2010; Zhong & Priest, 2011), including genetic stress (Tedman-Aucoin & Agrawal, 2012; Stevson, 2012; Heininger, 2013). Recombination may allow a population to keep up with environmental changes by producing appropriate novel allelic combinations (Robson et al., 1999; Manos et al., 2000; Carja et al., 2013). Work on the evolution of recombination rates in heterogeneous environments suggests that fluctuating selection may favor increased recombination when the direction of selection changes appropriately over time (Charlesworth, 1976, 1993; Lenormand & Otto, 2000; Otto & Michalakis, 2007).

5. Likewise, compelling evidence demonstrates that stress increases mutability of simple sequence repeats (SSRs) (Jackson, 1998; Wang et al., 1999, Nevo et al., 2005; Heininger, 2013), mobilization of transposable elements (Capy et al., 2000; Pericone et al., 2002; Heininger, 2013) and prion-driven phenotypic diversity (Haffmann et al., 2010). An immanent feature of SSRs is their high mutability, which leads to both sequence and length polymorphism (Kelkar et al., 2008; Pumpernik et al., 2008), the latter being at least one order of magnitude greater than the former (Borstnik & Pumpernik, 2002; Pumpernik et al., 2008). The SSR mutation rates ($10^{-6}$ to $10^{-9}$ events per locus per generation) are very high, as compared with the rates of point mutations at coding gene loci (Li et al., 2002; Ellegren, 2004). SSRs encode their own mutability through the unit size, length, and purity of the repeat tract (King et al., 1997; Ellegren, 2004; King & Kashi, 2007; Legendre et al., 2007). In 296 *Escherichia coli* genes related to repair, recombination and physiological adaptations to different stresses, Rocha et al. (2002) observed a significant high number of SSRs capable of inducing phenotypic variability by slipped-mispair during DNA, RNA or protein synthesis. Overrepresentation of SSRs in stress response genes may be a bacterial strategy to increase versatility under stressful conditions.

All these molecular processes jointly increase the stochasticity of the genotype-phenotype mapping under stressful conditions and variable environments (Altenberg, 1995; Wagner & Altenberg, 1996; El-Samad & Madhani, 2011; Martin et al., 2011). Overall, abiotic and biotic environmental conditions to which organisms are poorly adapted and that reduce Darwinian fitness elicit bet-hedging behavior increasing cellular noise, (epi)genetic and phenotypic diversity.

12. The gambles of life

Uncovering the mysteries of natural phenomena that were formerly someone else’s ‘noise’ is a recurring theme in science.

Bedar & Georges, 2000
12.1 Lottery and insurance: responses to uncertainty and risk

A growing consensus suggests that ecological and economic theories are ultimately indistinguishable (Boulding, 1978; Hirschleifer, 1977; Real & Caraco, 1986; Noé & Hammerstein, 1994; Gandolfi et al., 2002; Orr, 2007; Yaari & Solomon, 2010; Okasha & Binmore, 2012). Malthus’ (1803) argument that the growth rate of a population would tend to outpace the growth rate of output implied, for Darwin, an inevitable struggle for existence and, hence, natural selection of the fittest. Somewhat less well-known is the influence of Adam Smith (1776), whose “Invisible Hand” seems to have been a fundamental and pervasive inspiration for Darwin. Unfettered self-interested utility or profit maximization became, for Darwin, the struggle for reproductive success. The efficiency achieved by the market became the prodigious adaptation and balance evident in nature (Gould, 1993, p. 148–151; Robson, 2001).

There is a deep analogy between rational choice theory, particularly as it applies to games of strategy, and evolutionary theory. In a standard rational choice scenario, an agent is faced with a choice between a set of options; the aim of the theory is to say which choice is optimal. As Skyrms (1996, 2000) notes, it is easy to transpose such a scenario to an evolutionary context. Instead of thinking of a conscious agent trying to choose between the options, we can think of natural selection as doing the choosing, favoring the option with the greatest Darwinian fitness. Just as, according to traditional rational choice theory, the rational person favors the option that maximizes her/his expected utility, so natural selection favors the option that confers the greatest expected reproductive success on its bearer. Expected utility is thus analogous to expected number of offspring; the maximization of the former by rational agents is analogous to the maximization of the latter by natural selection. Just as rational choice theorists argue that much human behavior can be understood as an attempt to maximize expected utility, so evolutionary theorists argue that much animal behavior can be understood as an attempt to maximize reproductive output (Okasha, 2007).

Natural selection is often viewed as a statistical process, maximizing the expected or mean reproductive success of individuals carrying a certain gene or genotype (Darwin, 1859; Fisher, 1930). The expected reproductive success is then called ‘mean’ fitness. In this sense, standard theory can be referred to as a ‘statistical’ theory of natural selection. In order to analyze the optimality of a phenotypic trait based on mean fitness, most traditional theories of natural selection almost invariably assume constant and predictable environments. However, for almost all organisms in the wild, environments are variable and unpredictable (Yoshimura & Clark, 1993). In order to understand the basic properties of uncertainty, a probabilistic perspective for natural selection, a synthetic or integrated view of the effects of uncertainty on natural selection is warranted. Stochastic environments are the raffle boxes in the lotteries of life. Organisms have no choice other than to try their luck in these lotteries. On the other hand, insurance is the risk-sharing strategy of risk-averse agents that have to compete in lotteries.

It is a common observation that people exhibit risk-aversion when making some choices while also exhibiting risk-preference in other cases. People buy both insurance and lottery tickets. The standard explanation for this behavior begins with Friedman and Savage (1948), who suggested that the typical von Neumann-Morgenstern utility function is concave over low values of wealth but then becomes convex over higher values. People/animals with such utility/fitness functions would seek insurance protection against downside risk, while at the same time buying lottery tickets that promise a small probability of a large increase in wealth (Robson & Samuelson, 2009).

In economics, one way to take into account this effect was to declare that what is to be maximized is not the wealth itself but rather the “utility function” (von Neumann & Morgenstern, 1944). The case where the “utility function” is the logarithm of the wealth reduces to considering the geometric mean rather than the arithmetic mean. Thus, the use of this utility function may be interpreted as a way to take into account the fact that in general a strategy is applied repeatedly for long spans of time such that the frequency of the events approach their probability. Some of the behavioral anomalies studied over the years (Allais, 1953; Kahneman & Tversky, 1979; Thaler, 1994) can be related to the subtle difference between the expectation for one game and the probability for longer series of events (Yaari & Solomon, 2010).

12.2 “Decisions” under uncertainty: utility/fitness optimization

Yoshimura et al. (2009) classified environmental uncertainty into three categories based on the level of integration: (i) short-term temporal change experienced by an individual (individual level within a generation), (ii) phenotypic variation among individuals (population level within a generation) and (iii) population fluctuation across generations due to long-term environmental changes.
(cross-generation level). Knight (1921) made his famous distinction between risk and uncertainty by explaining that risk is ordinarily used in a loose way to refer to any sort of uncertainty viewed from the standpoint of the unfavorable contingency, and uncertainty similarly with reference to favorable outcomes. Uncertainty can be overcome by acquiring information about an environment (Stephens, 1987, 1989); risk cannot (Winterhalder et al., 1999). A large body of literature now has shown taxonomic ubiquitous risk-sensitive behavioral capacities (Stephens, 1981; Real & Caraco, 1986; Stephens & Krebs, 1986; Ellner & Real, 1989; Cartar & Dill, 1990; McNamara & Houston, 1992; Bernstein, 1996; Kacelnik & Bateson, 1996; Smallwood, 1996; Winterhalder et al., 1999; Dall & Johnstone, 2002; Shafir et al., 2005; Stephens et al., 2007; Ydenberg, 2007; Houston et al., 2011; Mayack & Naug, 2011; Ratikainen, 2012; Ito et al., 2013). When people and animals are faced with dicey decisions, a well-documented trend holds (Bernoulli, 1738/1954; Pratt, 1964; Arrow, 1965; Caraco & Chasin, 1984; Yoshimura & Shields, 1987; Hintze et al., 2013; Ito et al., 2013). If the stakes are sufficiently high, they are risk averse. Risk averseness is usually described as a resistance to accept a deal with risky payoff as opposed to one that is less risky or even safe, even when the expected value of the safer bargain is lower. The principle is similar to risk aversion in utility theory (Menezes & Hanson, 1970): the cost of a negative deviation from the mean is larger than the benefit of an equivalent positive deviation (Philippi & Seger, 1989).

Risk-sensitive behavior is variance-sensitive behavior (Smallwood, 1996; Ydenberg, 2007; Mayack & Naug, 2011; Ratikainen, 2012). The logic of risk-sensitive foraging is captured by the energy-budget rule of Stephens (1981). It can be illustrated using foragers facing two options with equal mean rewards but different variance (Rao & Sejnowski, 2003). Real and co-workers (Real et al., 1990; Real, 1991) performed a series of experiments on bumble bees foraging on artificial flowers whose colors, blue and yellow, predicted the delivery of nectar. They examined how bees respond to the mean and variability of this delivery in a foraging version of a stochastic two-armed-bandit problem. All the blue flowers contained 2 µl of nectar, 1/3 of the yellow flowers contained 6 µl, and the remaining 2/3 of the yellow flowers contained no nectar at all. In practice, 85% of the bees’ visits were to the constant-yield blue flowers despite the equivalent mean return from the more variable yellow flowers. When the contingencies for reward were reversed, the bees switched their preference for flower color within one to three visits to flowers. Real and co-workers further demonstrated that the bees could be induced to visit the variable and constant flowers with equal frequency if the mean reward from the variable flower type was made sufficiently high. This experimental finding shows that bumble bees, like honeybees, can learn to associate color with reward. Further, color and odor learning in honeybees has approximately the same time course as the shift in preference described above for bumble bees (Gould, 1987). It also indicates that under the conditions of a foraging task, bees prefer less variable rewards and compute the reward availability in the short term. This is a behavioral strategy used by a variety of animals under similar conditions for reward (Krebs et al., 1978; Real et al., 1990; Real, 1991), suggesting a common set of constraints in the underlying neural substrate.

The sensitivity to variance in food reward of small seed-eating birds (Caraco et al., 1980; Caraco, 1981, 1982, 1983; Caraco & Chasin, 1984; Caraco & Lima, 1985), shrews (Barnard & Brown, 1985), warblers (Moore & Simm, 1986), and hummingbirds (Stephens & Paton, 1986) is apparently affected by the probability of meeting daily energetic requirements. When the forager expects not to meet its energetic requirement, it should prefer the more uncertain alternative and hence be risk-prone. However, when it is doing well and expects not to fall short of its energetic requirement, it should avoid uncertainty and be risk-averse. Since stored reserves affect an organism’s expectation of meeting its food requirement, reserves should be a determinant of foraging risk-sensitivity. Thus, bumble bees can be both risk-averse (preferring constant flowers) and risk-prone (preferring variable flowers), depending on the status of their colony energy reserves. Diet choice in bumble bees appears to be sensitive to the “target value” of a colony-level energetic requirement (Cartar & Dill, 1990). Animals on a negative energy budget increase their preferences for risk, while animals on a positive energy budget are typically risk-averse (Rubenstein, 1982). Animals adapted to living in unpredictable conditions are unlikely to benefit from risk-seeking strategies, and instead are expected to reduce energetic demands while maintaining risk-aversion (Kahneman & Tversky, 1979; Stephens & Krebs, 1986; McNamara & Houston, 1992; Kacelnik & Bateson, 1996; Platt & Huettel, 2008; MacLean et al., 2012). However, if faced with a scenario in which the less variant food supply will not meet an animal’s expected energetic needs for survival, the animal should switch to a higher-risk food source that affords a greater chance of survival (Caraco et al., 1980; Caraco, 1981; Stephens, 1981). In other words, if the
rate of energy gain associated with the less variant, “safe” food supply falls short of that needed for survival, adopting a risk-seeking strategy offers the only chance of survival and should become the favored strategy. Because of their risk-taking attitude, lower-ranking individuals are more likely to innovate (Sigg, 1980; Katzir, 1983; Reader & Laland, 2001; Brosnan & Hopper, 2014). The malleability of these preferences may be evolutionarily advantageous, and important for maximizing chances of survival during brief periods of energetic stress (MacLean et al., 2012). The theory predicts, for example, that prey should select habitats that minimize the ratio of predation rate to growth rate (Werner & Gilliam, 1984). On the other hand, when food is scarce animals should increase their activity and use of risky habitats, thus increasing growth but also predation rates (Houston et al. 1993; Werner & Anholt, 1993; Mangel & Stamps, 2001). Studies in the laboratory (e.g. Gilliam & Fraser, 1987; Anholt & Werner, 1995, 1998) and in the field (Biro et al. 2003a, b) support these predictions and show substantial increases in prey activity, use of risky habitats and greater predation mortality with declines in food abundance (Biro et al., 2004). Predators select against high growth rates and risk-taking behavior in prey populations (Biro et al., 2004).

The fitness maximization problem arises when individuals have to choose among risky alternatives. Idiosyncratic risk, respectively uncertainty, is risk or uncertainty to which only specific agents are exposed, in contrast to systematic or aggregate risk/uncertainty that is faced by all agents in the market. For example, the weather is a standard example of aggregate risk—a very harsh winter may kill all members of a population. In evolution, often risk is a combination of a systemic component stemming primarily from the impact of unfavorable weather events and of an idiosyncratic component depending on individual characteristics. Lotteries may be either idiosyncratic or aggregate (or both) in nature. An idiosyncratic lottery is defined to be one where the realizations are statistically independent across individuals. A lottery is aggregate if individuals share outcomes. Curry (2001) showed that lotteries that involve idiosyncratic risk have differing implications for fitness from lotteries that involve aggregate uncertainty. This stems from the fact that nature selects for the gene with the highest growth rate within a population. When lotteries are purely idiosyncratic in nature, then reproductive value is equal to the individual's expected offspring. This is not true, however, for a lottery that has an aggregate component. An example given by Curry (2001) illustrated the difference between the two types of lotteries. The growth rate associated with an idiosyncratic lottery A is higher than that of an aggregate lottery B, even though B stochastically dominates A. It would seem that there would be strong selection for a gene that could make appropriate distinctions between lotteries that are idiosyncratic in nature and ones that involve an aggregate component.

Various theorists have addressed the role of variable environments, uncertainty and risk for the optimization of reproductive strategies. The common motif of these models is that individuals will randomize their strategies. Cooper and Kaplan (1982) have demonstrated that when lotteries are aggregate, the optimal decision rule involves randomization. That is, when the environment is stochastic, a gene may spread through the population faster when agents of the same genotype take different lotteries. This type of phenotypic variation was called "adaptive coin-flipping", "intra-genotypic strategy-mixing" (Cooper & Kaplan, 1982; Kaplan & Cooper, 1984; Cooper, 1989) or "stochastic polyphenism" (Walker, 1986). Strangely, Cooper and Kaplan (1982) termed this "coin-flipping altruism": "It is as though each individual of the superior strategy-mixing genotype were practising a form of "coin-flipping altruism" by assuming the risk of getting stuck with the personally inferior strategy... True, this is not the customary kind of altruism in which the altruist renders some tangible service to other individuals. It nonetheless represents a sacrifice of immediate individual fitness for the sake of the long term advantage of the genotype" (Cooper & Kaplan, 1982). According to this logic every participant of a lottery, by buying a lottery ticket, commits an act of altruism towards the eventual winner(s) of the lottery. Likewise, clients of insurance companies in which the insured event does not occur would act altruistically versus the clients in which the insured event occurs. And are gamblers at the casino and/or stock traders (Statman, 2002; Gao & Lin, 2011; Liao, 2013) when they lose altruists towards the winners?

With a fixed environment, the type of individual maximizing expected offspring is selected. In other words, the evolutionarily most successful attitude to risk is risk neutrality in offspring (Robson, 1996). Uncertainty due to a random environment has distinct evolutionary consequences from risk given the environment. The risks in the evolutionary environment are unlikely to have been purely idiosyncratic. Fluctuations in the weather or abundance of predators, epidemics, and failures of food sources are all bound to have a common effect on death rates. With a random environment, the type selected is strictly less averse to idiosyncratic risk than
to risk which is correlated across all individuals (Robson, 1996). This criterion incorporates an essential distinction between idiosyncratic risk given the environment and aggregate uncertainty concerning the environment itself, implying a greater aversion toward the latter. Insurance is a population-level risk-sharing strategy of risk-averse agents turning idiosyncratic risk into aggregate risk. Via the law of large numbers evolution generated a form of automatic biological insurance against idiosyncratic risk, whereas this insurance is inoperative in the same sense against aggregate uncertainty (Robson, 1996). The distribution of types of agents changes in response to generated rewards?this occurs through the standard replicator dynamic. In particular, preference types that do well, increase in relative frequency (To, 1999). The fundamental feature of multiplicative processes is the fact that the expected gain of the players taking part in this iterative process depends in a crucial way on the number of players considered (number of independent realizations) and the number of time steps that the game is played. For long times (the number of time steps played in the game), the expected wealth of the players follows the geometric mean and not the arithmetic mean of the game, keeping in mind that geometric mean ≤ arithmetical mean (Yaari & Solomon, 2010). The use of this function may be interpreted as a way to take into account the fact that in general a strategy is applied repeatedly for long spans of time such that the frequency of the events approach their probability (Yaari & Solomon, 2010).

13. Evolution is far-sighted

Evolution is not teleological in the sense that its processes or actions are for the sake of an end, i.e., the Greek “telo” or final cause. Clearly, once an organism has survived and/or reproduced one can point to its various attributes and say “yes, that attribute appears to have contributed to the organism’s survival/reproduction”. However, that is no more evidence of “foresightedness” than a lottery winner saying “I chose these lottery numbers (or bought those particular scratch-off tickets) because I knew they would be winners”. This is known as the “fallacy of affirming the consequent” (also called post hoc, ergo propter hoc argumentation) and is logically inadmissible in the natural sciences (MacNeill, 2009).

Backwards causation, by which some future state or event influences (‘causes’) an action in the present or past, is often characteristic of teleological arguments. The Modern Synthesis took pride in having discouraged such thinking (Mayr, 1992). In the tradition of the Modern Synthesis it has been argued that mutations must be random because natural selection cannot “assist the process of evolutionary change,” since “selection lacks foresight, and no one has described a plausible way to provide it” (Dickinson & Seger, 1999). Such an evolutionary strategy was called a raffle or lottery (Stockley et al., 1997; Parker et al., 2010) and would correspond to a “random trial” approach: genetic change would arise at random, independent of its functional consequences and natural selection would decide about its fitness value. However, even in a raffle competition an increased number of lottery tickets, as in sperm competition games (Parker, 1990), would increase the chances of a winner. Obviously, environments are uncertain and unpredictable. In consequence, evolution can have no foresight (Grant & Grant, 2002). However, like a chess player that takes several potential moves of his opponent into account, evolution is able to anticipate at least some of the more plausible “moves” of a stochastic environment and “takes them into account”, covering some of the more plausible bases.

The Baldwin effect, independently forwarded by Baldwin (1896), Lloyd Morgan (1896), and Osborn (1896), but largely so called because of Baldwin’s influential book (Baldwin, 1902), states that the ability of individuals to learn can guide and accelerate the evolutionary process (Hinton & Nowlan, 1987; French & Messinger, 1994; Weber & Depew, 2003; Sznajder et al., 2012; Weber, 2013). Currently, this principle is widely used in evolutionary computing and evolutionary algorithms (Ackley & Littman, 1991; Mitchell & Forrest, 1994; Bull, 1999; Eiben & Smith, 2008; Paenke et al., 2009a).

The Baldwin effect consists of the following two steps (Turney et al., 1996): In the first step, lifetime learning gives individual agents chances to change their phenotypes. If the learned traits are useful to agents and result in increased fitness, they will spread in the next population due to fitness-related differential reproduction. This step means the synergy between learning and evolution. In the second step, if the environment is sufficiently stable, the evolutionary path finds innate traits that can replace learned traits, because of the cost of learning. This step is known as genetic assimilation (Arita & Suzuki, 2000). Mathematical models suggest that learning would speed up the adaptation process by providing more explicit information about the environment in the genotype (Sendhoff & Kreutz, 1999; Arita & Suzuki, 2000). Learning alters the shape of the search space in which evolution operates and thereby provides good
evolutionary paths towards sets of co-adapted alleles. Hinton and Nowlan (1987) demonstrated that this effect allows learning organisms to evolve much faster than their non-learning counterparts, even though the characteristics acquired by the phenotype are not communicated to the genotype. With this feedback control adaptation proceeds by “trial and error” (Ashby, 1954). “Random trial” and “trial and error” approaches differ in an important variable: feedback of outcome. “Random trial” lacks the feedback loop: it either cannot find out whether the trial was a success or failure or it is completely unable to learn from this knowledge. The “trial and error” approach has a feedback loop that identifies “errors”. In evolution, the feedback loop occurs through Charles Darwin’s natural selection-mediated preferential reproduction of the fit (Ackley & Littman, 1991) and Alfred R. Wallace’s elimination of the unfit (Smith, 2012a, b). A learning system is able to draw its lesson from the error(s) and make its next trial less random. Learning from “trial and error” systems leads to “educated guess” approaches that are less random-driven but use past experience to navigate future direction and thereby limit the search space and increase the likelihood that some of the problem solutions generated will be useful (Jablonka & Lamb, 2007; Heininger, 2013). Learning may generate selection in favor of conspicuous novel traits faster, and for a wider range of traits, than genetically based sensory biases (Zuk et al., 2014).

Altenberg (2005) compiled a list of short-sighted adaptations:

- cheating, defection, and other antisocial behavior,
- meiotic drive (Lewontin, 1962),
- parthenogenesis (Griffiths & Butlin, 1995),
- overpopulation (Wynne-Edwards, 1962),
- imprudent predation (Rosenzweig, 1972) and other forms of habitat over-exploitation–the ‘tragedy of the commons’ (Hardin, 1968),
- cannibalism (Hamilton, 1970),
- cancer (the organism being the population) (Nunney, 1999; Stoler et al., 1999),
- adaptation to temporally unreliable resources (Kauffman & Johnsen, 1991),
- viable but infectious pathogen carrier states (Kirchner & Roy, 1999),
- evolution of endosymbionts to the detriment of host (Wallace, 1999).

He argued that when the trait confers short-term individual advantage but long-term population disadvantage, under a hierarchically structured population, evolvability evolves to be suppressed (Altenberg, 2005). Several processes have evolved that can tune the evolvability and far-sightedness of organisms:

(i) Condition-dependent mutagenesis
(ii) Epigenetic conditioning of mutations
(iii) Behavioral conditioning of new traits (Zuk et al., 2014).

A multitude of transgenerational processes indicate that evolution is far-sighted: evolution favors processes whose outcomes are robust and sustainable: in learning and memory past experience guides future actions (Kirschner & Gerhart, 2005; Gerhart & Kirschner, 2007; Parter et al., 2008); bet-hedging is a forward-looking response to past environmental unpredictability (Simons, 2009, 2011); demographic stochasticity and the tragedy of the commons is prevented by a multitude of processes establishing prudent reproduction (Goodnight et al., 2008, Heininger, 2013); evolution “cares” for future generations by curtailing the reproductive potential and lifespan of the current generation (Heininger, 2012); sexual reproduction is the paradigmatic bet-hedging process that creates pre-selected variation (Heininger, 2013).

14. Complexity and self-organization: chaos and order

...a fully adequate theory of evolution must encompass both self-organization and selection. Corning, 1995, p. 112

14.1 Complexity

To begin with, the term complex is a relative one. Individual organisms may use relatively simple behavioral rules to generate structures and patterns at the collective level that are relatively more complex than the components and processes from which they emerge. Systems are complex not because they involve many behavioral rules and large numbers of different components but because of the nature of the system’s global response. Complexity and complex systems generally refer to a system of interacting units that displays global properties not present at the lower level. These systems may show diverse responses that are often sensitively dependent on both the initial state of the system and nonlinear interactions among its components. Ever since the pioneering discovery by May (1974, 1976) in the seventies that simple rules can lead to complex dynamics including chaos,
ecological chaos has been a subject of intense research.

There are two quite different but complementary meanings of the term “complexity.” The term is used both to indicate randomness and structure. A correct understanding of complexity reveals that both are required elements of complex systems. A large number of cases demonstrate that structural complexity arises from the dynamical interplay of tendencies to order and tendencies to randomness (Crutcher & Machta, 2011). Complexity, in Ashby’s sense, is essentially conceived as a system’s potential to assume a large number of states, and we also have a measure for it: variety, the number of states a system can assume (Schwaninger, 2004).

The founding idea of complexity science was Prigogine’s juxtapositioning of the 1st and 2nd Laws of Thermodynamics so as to explain the emergence of dissipative structures (Stengers, 2004). Implicit in this was his questioning of the reversibility of time and the centrality of equilibrium in “normal” science (Prigogine & Stengers, 1984; Prigogine, 1997). Complexity science - really ‘order-creation science’ – is founded on theories explicitly aimed at explaining order creation rather than accounting for classical physicists’ traditional concerns about explaining equilibrium (McKelvey, 2001, 2004a, b). Systems that originate in response to, and are maintained by, the optimizing imperative of the 2nd Law of Thermodynamics are sometimes called dissipative structures (Nicolis & Prigogine, 1989) or complex adaptive systems (Levin, 1995). Prigogine (1997) asserted that like weather systems, organisms are unstable systems existing far from thermodynamic equilibrium. Instability resists standard deterministic explanation. Instead, due to sensitivity to initial conditions, unstable systems can only be explained statistically, that is, in terms of probability. Ilya Prigogine (1997) argued that complexity is non-deterministic, and there is no way whatsoever to precisely predict the future. There is the changing role of models. Math is good for equilibrium modeling. However, math models can’t handle order creation. Agent-based computational models are essential for modeling order creation (McKelvey, 2004a). The complex system approach is neither holistic nor reductionist but asserts that ecological relationships between patterns and processes span multiple scales of organization (Proulx, 2007). Many key concepts are often associated to the complex system approach: non-linearity, emergence, criticality, scaling, hierarchy and evolvability to list a few (Milne, 1998; Brown et al., 2002). On the other hand, the literature on nonlinear systems often mentions self-organization, emergent properties, and complexity as well as dissipative structures and chaos (Glansdor & Prigogine, 1971; Nicolis & Prigogine, 1989; Prigogine, 1997). Since these nonlinear interactions involve amplification or cooperativity, complex behaviors may emerge even though the system components may be similar and follow simple rules (Camazine et al., 2001). Emergent properties are features of a system that arise unexpectedly from interactions among the system’s components. An emergent property cannot be understood simply by examining in isolation the properties of the system’s components, but requires a consideration of the interactions among the system’s components (Kauffman, 1993; Kelso, 1995; Camazine et al., 2001; Corning, 2002). An ideal gas in a vessel of a macroscopic size is a large system because it contains 6 x 10^23 molecules per mole. This system, however, cannot be regarded as complex since all the elements interact by simple laws of classical or quantum mechanics that are uniformly applicable to all the events of interaction. One may call a system complex either if there is a wide variety of interactions between the system’s components, or if the system consists of a large number of distinctly different subsystems interacting with each other, or both (Rosenfeld, 2009). Complex biological systems manifest a large variety of emergent phenomena among which prominent roles belong to self-organization and swarm intelligence. On the other hand, emergence is what self-organizing processes produce (Corning, 2002). In fact, natural selection may well be an emergent phenomenon of the complex system “life” (Kauffman, 1993; Kelso, 1995; Weber & Depew, 1996; Weber, 1998; Hoelzer et al., 2006). Complex adaptive systems also require stochastic factors, e.g. noise and fluctuations (Gros, 2008). It is only with an intermediate level of stochastic variation, somewhere between determined rigidity and literal chaos that local interactions give rise to complexity (Johnson, 2001; Theise, 2004; Theise & Harris, 2006).

A complex system constantly changes, largely through three different types of transition (Manson, 2001): First, a key characteristic of a complex system is self-organization, the property that allows it to change its internal structure in order to better interact with its environment. Self-organization allows a system to learn through piecemeal changes in internal structure. Second, a system becomes dissipative when outside forces or internal perturbations drive it to a highly unorganized state before suddenly crossing into one with more organization (Schwie & Allen, 1982). Economies can be dissipative when confronting large shifts in the nature of their relationships with the environment. Introduction of new technologies, such
as in the industrial revolution, can spur radical change in the internal structure of an economy (Harvey & Reed, 1994). The work of Holling (1978, 1995) illustrates how small disturbances such as pest infestations or fire can trigger large-scale redistribution of resources and connectivity within the internal structure of an ecosystem.

Third, the term self-organized criticality refers to the ability of complex systems to balance between randomness and stasis. Instead of occasionally weathering a crisis, a system can reach a critical point where its internal structure lies on the brink of collapsing without actually doing so (Bak & Chen, 1991). Self-organized criticality is a form of self-organization where the rate of internal restructuring is almost too rapid for the system to accommodate but necessary for its eventual survival (Scheinkman & Woodford, 1994). Research on self-organized criticality is largely restricted to ecological and biogeophysical systems (e.g., Andrade et al., 1995; Correig et al., 1997) but there is a growing body of work on urban and economic systems (Sanders, 1996; Allen, 1997).

14.2 Fractals and 1/f noises

A particular class of complex systems are scale independent (Bak, 1996; Gisiger, 2001). A classical example of such systems in physics is the earth’s crust (Gutenberg, 1949; Turcotte, 1992). It is a well-established fact that a photograph of a geological feature, such as a rock or a landscape, is useless if it does not include an object that defines the scale: a coin, a person, trees, buildings, etc. This fact is described as scale invariance: a geological feature stays roughly the same as we look at it at larger or smaller scales. In other words, there are no patterns there that the eye can identify as having a typical size. The same patterns roughly repeat themselves on a whole range of scales. This property can manifest itself with fractals (spatial scale invariance), flicker noise or 1/f-noise where f denotes the frequency of a signal (temporal scale invariance) and power laws (scale invariance in the size and duration of events in the dynamics of the system). The patterns displayed by many natural systems do not allow for a simple description using Euclidean geometry: they present scale-invariance; that is, no characteristic length measure can be obtained from them. Therefore, when observed at different resolutions, they display the same pattern. The common feature of self-similar behavior is the presence of scaling laws (West et al., 1997; Gisiger, 2001) (also known as power laws). A wide variety of physical systems show power-law correlations in space (fractals) (Mandelbrot, 1982; Pietronero & Toscatti, 1986; Aharony & Feder, 1989) or time (1/f noises) (Voss, 1978; Weissman, 1988).

The Chaos Game shows that local randomness and global determination can coexist to create an orderly, self-similar structure called a fractal (Peters, 1994, p. 10–17). Fractals have the property of self-similarity in that the parts are in some way related to the whole. Fractal geometry is symbolized in the self-similar patterns of the Sierpinski triangle, which can be generated with an algorithm that has both a random and a lawful element (Carr, 2004). Natural beauty in mountains, plants, and snowflakes reveals a fractal geometry characterized by the complex interplay between randomness (symbolized by dice) and global determinism (which loads the dice) (Mandelbrot, 1983). Nature offers many examples of fractal statistics: branching in our lungs and in plants; variations in the flooding of the Nile river, of rainfall, of tree-rings and the intricate vein structure of leaves. Overall, evolution has a fractal geometry (Green, 1991; Burlando, 1993; Halley, 1996; Rikvold & Zia, 2003). The fluctuations of the stock market also obey fractal statistics (Peters, 1994). Because fractals involve long-range correlations, they also reflect some key features of how living systems are organized and how they evolve in time. The implications for evolution are very important, because cooperative effects emerging from the interactions can lead to new, sometimes counterintuitive, results. If fractal structures and self-similar fluctuations are so common, perhaps some universal dynamical processes are at work.

1/f noise represents the fluctuations of some physical quantity about its steady-state value. It is found in a wide variety of quite different systems (Voss & Clarke, 1978), and in some cases has been shown to be an equilibrium property (Voss & Clarke, 1976). Among ecologists, there has been a growing recognition of the importance of long-term correlations in environmental time series. Recent empirical evidence points to ever-increasing environmental variance through time (Steele, 1985; Pimm & Redfearn, 1988; Ario & Pimm, 1995; Bengtsson et al., 1997; Solé et al., 1997). The diversity of a desert ecosystem, for example, will be influenced by numerous small changes each day. Some rare events, such as desert storms, will have longer-lasting influence. The family of 1/f-noises – fluctuations defined in terms of the different timescales present – is a useful approach to this problem. White noise and the random walk, the two currently favored descriptions of environmental fluctuations, lie at extreme ends of this family of processes. A true random process or white noise has no correlations in time. Recent analyses of data, results of models, and
examination of basic $1/f$-noise properties, suggest that pink $1/f$-noise, which lies midway between white noise and the random walk, might be the best null model of environmental variation (Halley, 1996). There is strong evidence that background abiotic fluctuations have $1/f$-noise spectra (Mandelbrot & Wallis, 1969; Steele, 1985), though there may be significant differences between terrestrial and marine environments (Steele, 1985; Vasseur & Yodzis, 2004). A white noise represents the maximum rate of information transfer, but a $1/f$ noise, with its scale-independent correlations, seems to offer the best compromise between efficient information transfer and immunity to errors on all scales (Voss, 1992). Spectral density measurements of individual DNA base positions suggest the ubiquity of low-frequency $1/f$ noise and long-range fractal correlations as well as prominent short-range periodicities (Voss, 1992). Lévy flights, a special class of Markov processes, are scale invariant and often associated with power-laws described in many systems (Cole, 1995; Viswanathan et al., 1999; Martin et al., 2001). Lévy-like search strategies were revealed in analyses of a variety of behaviors from plankton to humans (Berg, 1993; Viswanathan et al., 1996, 2001; Bartumeus et al., 2003; Barabasi, 2005; Brockmann et al., 2006; Reynolds & Frye, 2007; Reynolds & Rhodes, 2009; Humphries et al., 2010). The models simulating these behaviors combine a multitude of stochastic processes by deterministic rules (Maye et al., 2007). In addition to the inevitable noise component, a nonlinear signature suggesting deterministic endogenous processes (i.e., an initiator) is involved in generating behavioral variability. It is this combination of chance and necessity that renders individual behavior so notoriously unpredictable (Maye et al., 2007).

14.3 Self-organization

A basic feature of diverse systems is the means by which they acquire their order and structure (Camazine et al., 2001; Ben-Jacob, 2003). In self-organizing systems, pattern formation occurs through interactions internal to the system, without intervention by external directing influences. Haken (1977, p. 191) illustrated this crucial distinction with an example based on human activity: “Consider, for example, a group of workers. We then speak of organization or, more exactly, of organized behavior if each worker acts in a well-defined way on given external orders, i.e., by the boss. We would call the same process as being self-organized if there are no external orders given but the workers work together by some kind of mutual understanding.” (Because the “boss” does not contribute directly to the pattern formation, it is considered external to the system that actually builds the pattern.) Systems lacking self-organization can have order imposed on them in many different ways, not only through instructions from a supervisory leader but also through various directives such as blueprints or recipes, or through pre-existing patterns in the environment (templates).

Critical to understanding Camazine’s et al. (2001) definition of self-organization is the meaning of the term pattern. As used here, pattern is a particular, organized arrangement of objects in space or time. Examples of biological pattern include a school of fish, a raiding column of army ants, the synchronous flashing of fireflies, and the complex architecture of a termite mound. To understand how such patterns are built, it is important to note that in some cases the building blocks are living units—fish, ants, nerve cells, etc.—and in others they are inanimate objects such as bits of dirt and fecal cement that make up the termite mound. In each case, however, a system of living cells or organisms builds a pattern and succeeds in doing so with no external directing influence, such as a template in the environment or directions from a leader. Instead, the system’s components interact to produce the pattern, and these interactions are based on local, not global, information. In a school of fish, for instance, each individual bases its behavior on its perception of the position and velocity of its nearest neighbors, rather than knowledge of the global behavior of the whole school. Similarly, an army ant within a raiding column bases its activity on local concentrations of pheromone laid down by other ants rather than on a global overview of the pattern of the raid (Camazine et al., 2001).

It seems that the philosopher Kant was the first to define life as a “self-organized, self-reproducing” process (Karsenti, 2008). Through pure reasoning, he defined life as the emergence of functions by self-organization. He said that in an organism, every part owes its existence and origin to that of the other parts, with the functions that are attributed to a complete living organ or organism emerging from the properties of the parts and of the whole. He defined this complex state of living matter as a self-organized end (Kant, 1790; Van de Vijver, 2006; Fox Keller, 2007). This led him to question the validity of using the causality principle of classical physics to explain life, and to suggest that a new kind of science would be required to study how purpose and means are intricately connected (Kant, 1790).

The cybernetician W. Ross Ashby (1962) proposed what he called “the principle of self-organization”. He noted that a dynamical system, independently of its type or composition, always tends to evolve towards a
state of equilibrium, or what would now be called an attractor. This reduces the uncertainty we have about the system's state, and therefore the system's statistical entropy. This is equivalent to self-organization (Heylighen, 2001). The resulting equilibrium can be interpreted as a state where the different parts of the system are mutually adapted. Heinz von Foerster (1960) formulated the principle of "order from noise". In a similar vein of thought, self-organization was proposed to generate "order through fluctuations" (Prigogine & Stengers, 1984) or "order through disorder" (Saetzler et al., 2011). Von Foerster noted that, paradoxically, the larger the random perturbations ("noise") that affect a system, the more quickly it will self-organize (produce "order"). Another reason for this intrinsic robustness is that self-organization thrives on randomness, fluctuations or "noise". In fact, self-organizational phenomena depend deeply on stochastic processes (Weber & Depew, 1996). The polarity of randomness and law characterizes the self-creating natural world (Carr, 2004). Non-linear systems have in general several attractors. When a system resides in between attractors, it will be in general a chance variation, called "fluctuation" in thermodynamics, that will push it either into the one or the other of the attractors. Without the initial random movements, the spins would never have discovered an aligned configuration. It is this intrinsic variability or diversity that makes self-organization possible. Just as there is an optimal level of stochastic perturbation favoring adaptive responses of thermodynamic self-organization systems (Helbing & Vicsek, 1999), there is an optimal rate of mutation that favors adaptive evolution under natural selection (Iwasaki & Yonezawa, 1999).

A logical consequence of complexity science as 'order-creation science' (McKelvey, 2004a, b) is that order could be generated through evolution by a synergy between natural selection and self-organizing processes (Solé et al., 1999). A diverse group of researchers in mathematics, physics, and several branches of biology have argued that self-organization should be placed alongside natural selection as a complementary mechanism of evolution (Nicolis & Prigogine, 1977; Conrad, 1983; Kaufman, 1993, 1995; Corning, 1995; Camazine et al., 2001; Heylighen, 2001; Richardson, 2001; Denton et al., 2003; Kurakin, 2005b, 2007; Hoelzer et al., 2006; Newman et al., 2006; Karsenti, 2008; Wills, 2009). Using computer models, Michael Conrad (1983) showed that natural selection will work together with self-organization to "smooth out fitness landscapes", thereby reducing the large differences in a Wrightian fitness landscape to shallow saddles. In this way self-organization plays a role even on highly gradualistic, Fisher-like assumptions. Natural selection, and the adaptations it brings about, necessarily occur within a veritable ocean of stochastic and self-organizational events and processes. Stochastic, selective, and self-organizational processes are empirically and causally intertwined in the evolution of living systems (Weber & Depew, 1996). Self-organization fails to emerge in completely determined systems (planets in motion, billiard balls) and completely random ones (molecules in a gas). It is only with an intermediate level of stochastic variation, somewhere between determined rigidity and literal chaos that local interactions give rise to complexity (Johnson, 2001; Theise, 2004; Theise & Harris, 2006). It is at the boundary between order and chaos at which evolvability is maximized. The highly ordered regime is one in which perturbations generate minimal overall change. In other words, there is minimal variation. The chaotic regime is one in which perturbations have unpredictable effects: there is no correlation between the initial and perturbed states. There is no heritability. The "edge of chaos" that is, in the narrow domain between frozen constancy (equilibrium) and turbulent, chaotic activity, is simply the region in which there is heritable variation. Heritable variation is necessary for evolution. Systems at the "edge of chaos" would be "evolvable," because it is only here that there is heritable variation (Kauffman, 1993, 1995; Richardson, 2001).

### 14.3.1 Self-organized criticality

It is becoming increasingly clear that many complex systems have critical thresholds—so-called tipping points—at which the system shifts abruptly from one state to another (Scheffer et al., 2009). In medicine, spontaneous systemic failures such as asthma attacks (Venegas et al., 2005) or epileptic seizures (Litt et al., 2001; McSharry et al., 2003) occur; in global finance, there is concern about systemic market crashes (Kambhu et al., 2007; May et al., 2008); in the Earth system, abrupt shifts in ocean circulation or climate may occur (Lenton et al., 2008); and catastrophic shifts in rangelands, fish populations or wildlife populations may threaten ecosystem services (Scheffer et al., 2001; Millennium Ecosystem Assessment, 2005). Cooperation is an emergent behavior of complex systems (Miramontes & DeSouza, 2014; Heininger, 2015): under adverse environmental conditions unicellular microorganisms display multicellular behavior (Juhas et al., 2005; Hooshanghi & Bentley, 2008). The mechanism by which complex systems tend to maintain on this critical edge has been called...
self-organized criticality (Bak et al., 1987; Bak, 1996; Jensen, 1998). The system’s behavior on this edge is typically governed by a “power law”: large adjustments are possible, but are much less probable than small adjustments. Self-organized criticality (SOC) is stated as follows: large, far from equilibrium, complex systems, formed by many interacting parts, spontaneously evolve towards the critical point. SOC was originally introduced (Bak et al., 1987) as an approach to understand 1/f-noise as well as the apparent abundance of power laws in nature, which is generally accepted as the sign of scale-invariance. The idea is that under very general circumstances driven stochastic processes develop into a scale-invariant state without the explicit tuning of parameters, contrary to what one would expect from equilibrium critical phenomena (Stanley, 1971; Pruessner, 2004). SOC systems can self-tune to a balanced (critical) state, precisely at the transition between a (subcritical) regime of inactivity and one of (supercritical) runaway activity. Sightings of SOC have been reported in every conceivable and inconceivable area of science (Clauset et al., 2009), including sociology (Roberts & Turcotte, 1998; Bentley & Maschner, 2000), financial markets (Lux & Marchesi, 1999), computer science (Gorshenev & Pis’mak, 2004; Cook et al., 2005), computer network traffic (Fukuda et al., 2000; Valverde & Solé, 2002), engineering (Carreras et al., 2004), and biology (Bak & Sneppen, 1993; Sepkoski, 1993; Sneppen et al., 1995; Solé et al., 1999; Bornholdt & Rohlf, 2000; Camazine et al., 2001; Nykter et al., 2008, Ribeiro et al., 2010; Mora & Bialek, 2011; Furuwasa & Kaneko, 2012; Longo et al., 2012; Krotov et al., 2014).

A simple metaphor of an SOC process is provided by a sandpile (Bak et al., 1987; Bak, 1996). If additional sand grains are randomly added on top of a sand pile then inevitably an instance will occur when local steepness of the slope surpasses a certain critical threshold thus causing local failure of structural stability. The excess of material will cascade into adjacent areas of the pile causing their failures as well. Thus an avalanche will occur, shifting the entire sandpile into a new stable state. What is fundamentally important in this process is that a random local event quickly propagates through the entire system, thus establishing long-range correlations within the system. SOC exemplifies an emergent phenomenon of system-wide organized behavior resulting from purely mechanistic reasons, i.e. from member-to-member local interactions without any intelligent organizing force (Rosenfeld, 2013). This new state cannot be anticipated from the properties of individual units.

In physics, fractal structures in space and time are known to emerge in the proximity of some types of phase transition (Binney et al., 1992; Solé et al., 1996b). The classic example is a magnetic material. A small piece of iron can tug on a paper clip at room temperature, but when heated to a high temperature no magnetic power is observed. The atoms that form the iron are themselves like small magnets. Each atom only interacts with its nearest neighbors and their natural tendency is to align spontaneously into small domains with the same orientation. At high temperature the coupling between nearest atoms breaks down because of thermal perturbations and, therefore, the atoms can have any polarity (up or down). But suddenly, when the material is cooled down, order spontaneously shows up. There is a critical temperature at which global magnetization appears and both fractal-spatial and fractal-temporal features arise. These transitions are described by an ‘order parameter’ which is zero at the disordered phase and positive otherwise (Solé et al., 1999).

The hypothesis that tuning a biological system to a critical state would render it somehow optimal has a long history (Langton, 1990). The underlying idea is that a system tuned to criticality presents a richer dynamical repertoire, being therefore able to react (i.e. process information) to a wider range of challenges (environmental or other) (Ribeiro et al., 2010). The experimental evidence in this direction ranges from gene expression patterns in response to stimulation of single macrophages (Nykter et al., 2008) to collective ant foraging (Beekman et al., 2001). Gene regulatory networks operate in a critical regime, i.e. close to a phase transition between ordered and chaotic dynamics (Serra et al., 2004, 2007; Shmulevich et al., 2005; Balleza et al., 2008; Nykter et al., 2008; Chowdhury et al., 2010; Torres-Sosa et al., 2012). Criticality is profoundly linked to evolvability. Critical dynamics, and hence the developmental trade-off in genetic networks, naturally emerge as a robust byproduct of the evolutionary processes that select for evolvability and optimize the evolutionary trade-off. Furthermore, the emergence of criticality occurs without fine-tuning of parameters or imposing explicit selection criteria regarding specific network properties (Torres-Sosa et al., 2012). Complex adaptive or evolutionary systems can be much more efficient in coping with diverse heterogeneous environmental conditions when operating at criticality. Analytical as well as computational evolutionary and adaptive models vividly illustrate that a community of such systems dynamically self-tune close to a critical state as the complexity of the environment increases while they remain noncritical for simple and
The capability to perform complex computations, which turns out to be the fingerprint of living systems, is enhanced in "machines" operating near a critical point (Langton, 1990; Kauffman, 1993; Bertschinger & Natschläger, 2004), i.e. at the border between two distinct phases: a disordered phase, in which perturbations and noise propagate unboundedly—thereby corrupting information transmission and storage—and an ordered phase where changes are rapidly erased, hindering flexibility and plasticity. The marginal, critical situation provides a delicate compromise between these two impractical tendencies, an excellent tradeoff between reproducibility and flexibility (Beggs, 2008; Chialvo, 2010; Shew & Plenz, 2013) and, on larger time scales, between robustness and evolvability (Wagner, 2005, 2008).

At the molecular level, an example of a critical point in biological systems can be exemplified by the dynamics of RNA viruses (Solé et al., 1999; Domingo et al., 2001, 2012). The understanding of how these entities evolve and adapt must take into account their extreme variability, caused by the error-prone RNA polymerase activity and the lack of proofreading mechanisms (Nowak, 1992; Heininger, 2013). Instead of a given single sequence, there is a cloud of mutants around the so-called ‘master’ sequence. This cloud is known as a quasispecies, a term first coined by Manfred Eigen (Eigen, 1971; Eigen et al., 1988). RNA viruses adapt to a changing environment by making use of their variability. Selection pressures by the immune system force the virus quasispecies to evolve (Kamp et al., 2003). The quasispecies model is consistent with this observation, but something defeats our intuition: there is a critical mutation rate beyond which heredity breaks down. This is referred to as the error catastrophe, and it is nothing but a phase transition point, which poses serious limitations to viral complexity. Available molecular data confirm the theory: RNA viruses do replicate close to the error catastrophe (Swetina & Schuster, 1982; Schuster, 1994; Cottry et al., 2001; Anderson et al., 2004). Eigen (1992) argued that virus replication error rates established themselves near an error-threshold where the best conditions for evolution exist.

Sexual reproduction displays another SOC phenomenon. Sexual reproduction uses a variety of stressors to create variation and to select the most resilient gametes and offspring from this variation (Heininger, 2013). Oxidative stress is an inherent feature of gametogenesis in all taxa. Importantly, from lower to higher taxa there is a substantially incremental use of this general principle. As evidenced by a male mutagenic bias, particularly male gametogenesis balances at the verge of mutational error catastrophe. The phenotype of the transition to error catastrophe is characterized by infertility.

In the light of the SOC theory, the variation-creating processes (discussed in chapters 11.1 and 11.2) and their tuning under stress (see chapter 11.3) can be interpreted as selected-for phenomena in self-organizing systems at the threshold of criticality.

15. Stochasticity and multilevel selection

...that an opinion has been widely held is no evidence whatever that it is not utterly absurd....
Bertrand Russell (1929)

In what follows, ‘individual’ refers to an individual organism, whereas a population refers to ‘a group of conspecific organisms that occupy a more or less well-defined geographic region and exhibit reproductive continuity from generation to generation’ (Futuyma [1986], pp. 554–5). Even though a population is composed of individual organisms, it is important to distinguish between properties that apply to individual organisms and properties that characterize the relationships among organisms—that is, properties that apply to populations. For example, individual organisms have properties such as color, shape and length. Populations, on the other hand, have properties such as size (defined as the number of individuals), frequency (defined as the proportion of individuals of one type or another) and growth rate (defined as the rate of change in the number of individuals in the population). Thus, in a sense, population-level properties are properties that arise only given the collection and interaction of individuals (Millstein, 2006). There are a variety of evolutionary phenomena that depend on population-level processes. Moreover, diversity and variation are population-level properties. Frequency- and density-dependent selection depend on population-level selection because the outcome of selection (the change in gene or genotype frequencies from one generation to the next) are determined by population-level parameters: the frequency of genotypes within a population or the density of the population (Millstein, 2006). Stochastic environments change the rules of evolution. Lotteries cannot be played and insurance strategies not employed with single individuals. These are emergent population-level processes that exert population-level
selection pressures generating variation and diversity at all levels of biological organization. Together with frequency and density-dependent selection, lottery- and insurance-dependent selection act on population-level traits.

Recent discussions by philosophers of science regarding natural selection have given conflicting answers to a pair of questions: first, is natural selection a causal process or is it a purely statistical aggregation? And second, is natural selection at the population level or at the level of individuals? Walsh et al. (2002) and Matthen & Ariew (2002) argued that natural selection is purely statistical and on the population level, whereas Bouchard & Rosenberg (2004) maintained that natural selection is causal and on the individual level. Millstein (2006) argued for a third possibility: natural selection is indeed a causal process, but it operates at the population level. What causes the conceptual confusion is the feedback control of the cybernetic system (figure 1C). This involves the replacement of the open, linear, chain of cause and effect familiar in most science by a circular causality, a closed feedback loop that implies the merging of causes and effects, the confluence of output and input signals. The system, however, cannot be understood properly without conceptually distinguishing input and output signals. Importantly, the output signal is a population-level signal including density- and frequency-dependent selection act on

frequency and density-dependent selection, lottery-

and insurance-dependent selection act on population-level traits.

A very large proportion of the literature pertaining to group selection consists of theoretical papers. The general conclusion has been that, although group selection is possible, it cannot override the effects of individual selection within populations except for a highly restricted set of parameter values. Since it is unlikely that conditions in natural populations would fall within the bounds imposed by the models, group selection, by and large, has been considered an insignificant force for evolutionary change (Wade, 1978a). David Sloan Wilson, Elliott Sober and a growing number of other workers has been attempting to resurrect group selection on a new foundation. What Wilson calls “trait group selection” (Wilson 1975, 1980; Wilson & Sober 1989, 1994) refers to a model in which there may be linkages (a “shared fate”) between two or more individuals (genotypes) in a randomly breeding population, such that the linkage between the two becomes a unit of differential survival and reproduction (Corning, 1997). Compatible with this concept, in game theory being applicable to fluctuating environments the players need never physically interact, compete or even communicate (Hutchinson, 1996). John Maynard Smith (1982b) developed a similar model, which he dubbed “synergistic selection”, in recognition of the fact that it implies a functional interdependency. According to Corning (1997), functional synergy explains the evolution of cooperation in nature, not the other way around. In other words, functional groups (in the sense of functionally integrated “teams” of cooperators of various kinds) have been important units of evolutionary change at all levels of biological organization; “functional group selection” is thus a ubiquitous aspect of the evolutionary process (Corning, 1997). George Price (1970, 1972) provided an elegant formalization that showed, among other things, how
the force of natural selection acting on genes can be partitioned into ‘group-level’ and ‘individual-level’ components. Unfortunately, the insight derived from Price’s simple demonstration did not spread very far outside of theoretical evolutionary biology and failed to impede the spread of the belief that group-selectionist-thinking is somehow logically flawed, wrong-headed, or merely wishful thinking. This untutored dismissal of group selection has slowed progress in understanding a variety of evolutionary processes (Henrich, 2004). However, with the exception of some orthodox Darwinists (e.g. Dawkins, 2012), multilevel selection (i.e., individual and group selection combined) has now received broad support (e.g. Gould, 2002; Okasha, 2006; Bijma et al., 2007; Godfrey-Smith, 2009; Calcott & Sterelny, 2011; Nowak & Highfield, 2011; Edward O. Wilson, 2012).

Maynard Smith (1976) argued that “For group selection, the division into groups which are partially isolated from one another is an essential feature. [...] that the extinction of some groups and the ‘reproduction’ of others are essential features of evolution by group selection. If groups are the units of selection, then they must have the properties of variation, multiplication, and heredity required if natural selection is to operate on them.” Early theorists (e.g. Williams, 1966; Maynard Smith, 1976) made evolutionarily unrealistic assumptions about group selection. In the light of empirical studies of group selection with laboratory populations of the flour beetle, Tribolium (Wade, 1976, 1977), Wade (1978a) argued that the models have a number of assumptions in common which are inherently unfavorable to the operation of group selection. (Keep in mind: “A mathematical model is only as good as its assumptions.” [Maynard Smith & Brookfield, 1983]). In their group selection experiments with Impatiens capensis, Stevens et al. (1995) showed that groups need not be discrete entities (Goodnight, 2005). Rather, groups were defined by interactions and their effect on fitness (Stevens et al., 1995). Within this framework, coevolution, e.g. of host-parasite and prey-predator communities (Gilpin, 1975; Levin & Pimentel, 1981), convergent evolution (Orians & Paine, 1983), the guild concept (Simberloff & Dayan, 1991), and community and ecosystem phenotypes (Whitham et al., 2003, 2006) are no longer conceptual orphans.

One of the cases, illustrating the biased choice of model assumptions, concerns the treatise of bet-hedging by theoreticians. There are two distinct forms of bet-hedging: (i) between-generation and (ii) within-generation. Current theory predicts that bet-hedging is far more likely to be a successful evolutionary strategy when the bets are hedged over several generations, than in a within-generation scenario (Yasui, 1998; Hopper et al., 2003). According to theory, the time scale of between-generation bet-hedging ensures that all individuals with a given phenotype suffer the same fate—circumstances such as drought exert homogenous pressure on all members of a population. Under within-generation bet-hedging, however, individuals with the same phenotype are subject to heterogeneous selection pressure—for example, will affect some individuals but not others. An important consequence of this difference is that conditions favoring the evolution of within-generation bet-hedging are very restricted. While a single lineage may realize increased fitness via within-generation bet-hedging, this fitness advantage varies inversely with population size and becomes vanishingly small at even modest population sizes (Yasui, 1998; Hopper et al., 2003). Most students of evolution are trained to focus on costs and benefits at the individual level, and tend to seek adaptive explanations for individual traits such as bet-hedging (e.g., Grafen, 1999). Although this focus is often successful, it leads astray in the case of within-generation bet-hedging. Only by assessing the fitness effects of a trait in the context of whole populations can one accurately identify traits that can and cannot be favored by within-generation bet-hedging (Hopper et al., 2003). Polyandry (Yasui, 1998; see chapter 16.3), conspecific brood parasitism (see below), prolonged diapause (Menu & Desouhant, 2002; see chapter 15.3.1.4) or cooperation (Fronhofer et al., 2011; Rubenstein, 2011) are within-generation bet-hedging phenotypes that have a much higher prevalence than theory would predict.

In the case of conspecific brood parasitism, a within-generation bet-hedging behavior, the biased choice of model assumptions has been refuted by long-term field data (Pöysä & Pesonen, 2007). Conspecific brood parasitism (CBP) is a taxonomically widespread alternative reproductive tactic in which a female lays eggs in the nest or egg group of a conspecific that provides all subsequent parental care (de Valpine & Eadie, 2008; Lyon & Eadie, 2008). CBP is particularly widespread among birds, being documented in at least 234 species and is particularly prevalent in Anseriformes where it has been reported in 76 of the 161 species (Payne, 1977; Yom-Tov, 2000, 2001). Moreover, it occurs in several other animal taxa, including fishes (e.g., Sato, 1986; Wisenden, 1999), amphibians (e.g., Summers & Amos, 1997), and insects (e.g., Eckwert, 1975; Eberhard, 1986; Müller et al., 1990; Zink, 2000, 2003; García-González & Gomendio, 2003; Loeb, 2003; Tallamy, 2005).
Because con specifics provide the only hosts for brood parasites, obligate parasitism cannot become fixed in a population. Further, the advantages of parasitic laying are likely to be greatest when the frequency of parasitism is low and many host nests are available containing few parasitic eggs; the advantages will decrease as frequency of parasitism increases and more host nests contain many parasitic eggs (de Valpine & Eadie, 2008). One of the earliest hypotheses to explain the occurrence and evolution of CBP was that by spreading eggs among nests, parasites can increase the likelihood that at least some offspring will escape predation and survive to independence, also known as the “risk spreading” hypothesis (e.g., Rubenstein, 1982; Petrie & Møller, 1991). Specifically, on the basis of a simulation model, Rubenstein (1982) reached the conclusion that laying eggs in several nests to avoid predation has a selective advantage over laying all the eggs in one nest. Indeed, considering that nest predation is the major source of nesting mortality in birds (Rickles, 1969; Nilsson, 1984; Martin 1988; Wesolowski and Tomialoj?, 2005) and plays an important role in the life-history evolution of birds (Bosque & Bosque, 1995; Martin 1995; Martin & Clobert, 1996; Saether, 1996; Julliard et al., 1997; Martin et al., 2000; Ghalambor & Martin, 2001), risk spreading is an appealing explanation for the evolution and occurrence of CBP (Pöysä & Pesonen, 2007). However, assuming that nests are predated at random and that parasites lay eggs randomly with respect to nest predation risk, Bulmer (1984) found that different egg-distribution strategies produced the same mean fitness when entire clutches were affected by stochastic events. Empirical field work in a well-studied model species of CBP, the common goldeneye (Bucephala clangula), revealed that, at least in some species, these assumptions are not valid. Pöysä and coworkers (Pöysä 1999, 2003, 2006; Pöysä et al., 2001, 2014) found that nests are not predated at random and that parasites use risk assessment and preferentially lay in safe nests. By taking these findings into account, model simulations revealed that the selective advantage of parasitic egg laying related to nest predation is much higher than previously thought (Pöysä & Pesonen, 2007). Likewise, by modeling mean fitness under a variety of egg-distribution strategies with only partial nest predation (as is often observed in nature), Roy Nielsen et al. (2008) found that higher fitness resulted from distributing eggs among multiple nests.

Cooperation is also a within-generation bet-hedging response that can be both conservative and diversifying (Fronhofer et al., 2011; Rubenstein, 2011). Typically, individual fitness and population fitness are in conflict. While selfish behavior is favored by individual selection, cooperation can evolve in many models of multilevel/group selection (Eshel, 1972; Uyenoyama, 1979; Slatkin, 1981; Leigh, 1983; Wilson, 1983; Boyd & Richerson, 1990, 1992, 2002; Binmore, 1992, 1994a, b; van Baalen & Rand, 1998; Bergstrom, 2002; Goodnight, 2005; Killingback et al., 2006; Traulsen & Nowak, 2006; Nowak et al., 2010; see Heininger, 2015).

With Starrfelt and Kokko (2012) I think that the distinction between within- and between-generation bet-hedging is flawed. Ignoring such artificial distinctions, a recent model (Ratcliffe et al., 2015) examined how key characteristics of risk and organismal ecology affect the fitness consequences of variation in diversification rate. In 1000-patch metapopulations the spatial and temporal dynamics of uncertainty were modeled. Either small (10 individuals) or large (100 individuals) carrying capacities, resulted in maximum global population sizes of 10^2 or 10^3 individuals, respectively. A single unpredictable event varied in scale from population-wide (e.g., a landscape-level process like unpredictable season length) to local (e.g., chance of nest discovery by a predator). Similarly, risk affected populations randomly in time or occurred in correlated series. Rapid diversification was strongly favored when the risk faced has a wide spatial extent, with a single disaster affecting a large fraction of the population. This effect was especially great in small populations subject to frequent disaster. In contrast, when risk was correlated through time, slow diversification was favored because it allows adaptive tracking of disasters that tend to occur in series. Naturally evolved diversification mechanisms in diverse organisms facing a broad array of environmental risks largely supported these results. The theory explained the prevalence of slow stochastic switching among microbes and rapid, within-clutch diversification strategies among plants and animals (Ratcliffe et al., 2015).

In contrast to what theoretical models suggest, group selection concepts have strong empirical support. When resources are limited, adult productivity in experimental populations of the flour beetle Tribolium was found to be strongly and negatively correlated with time to extinction of populations (MacDonald & Stoner, 1968, Nathanson, 1975; Wade, 1977). Thus, individual fitness, measured as relative reproductive rate, and population fitness, measured as persistence, may be in conflict. Wynne-Edwards (1962) discussed this possibility in detail and suggested that interpopulation selection may have led to the evolution...
of controls on individual reproductive interests. Multilevel selection analyses find that sizes appear as a competitive/selfish trait, favored in individual selection but selected against in group selection (Stevens et al., 1995; Aspi et al., 2003; Donohue, 2004; Weinig et al., 2007; Boege, 2010; Dudley et al., 2013). Plant height or elongation often appear as competitive traits in multilevel selection studies; they were selected to increase by individual selection and decrease by group selection in four studies (Stevens et al., 1995; Donohue, 2003, 2004; Weinig et al., 2007), selected to increase by individual and group selection in Silene (Aspi et al., 2003) and increase under individual selection only in another study (Boege, 2010). These opposing forces of selection result in an ecological process frequently observed in plants: a constant seed yield regardless of planting density (Goodnight et al., 1992; Stevens et al., 1995; Goodnight & Stevens, 1997; Donohue, 2003; Weinig et al., 2007). More specifically, individual selection can reasonably be expected to prevail in lower-density stands and favor large individual size, while selection at the group level may predominate in higher-density stands and act to reduce individual size (Weinig et al., 2007). In Tribolium, the ecological mechanisms of interspecies competition are the same, for the most part, as those of intraspecific competition (Park et al., 1964, 1965, 1974, Teleky 1980). The evolutionary interests of individuals within populations may be different from, and possibly opposed to, the evolutionary interests of populations (Wade, 1980a; Goodnight, 1985). Empirical studies have confirmed that group selection can be effective in situations when individual selection is not (Craig, 1982; Goodnight, 1985, 1990) and leads to faster evolutionary change than individual selection alone (Wade, 2003). Genetically-based interactions between individuals will not respond to individual selection but will respond to group selection (Griffing, 1977; 1981a, b). These findings support Wade’s (1978) suggestion that higher level selection can act on sources of genetic variance that is not available to lower levels.

15.1 Community selection as an emergent behavior of complex systems
As discussed previously (see chapter 14.1) complexity and complex systems generally refer to a system of interacting units that display global properties not present at the lower level. In their group selection experiments with Impatiens capensis, Stevens et al. (1995) showed that groups need not be discrete entities (Goodnight, 2005). Rather, groups are defined by interactions and their effect on fitness (Stevens et al., 1995). Although quantitative genetics has successfully been applied to many traits, it does not provide a general theory accounting for interaction among individuals and selection acting on multiple levels. Consequently, current quantitative genetic theory fails to explain why some traits do not respond to selection among individuals, but respond greatly to selection among groups (Bijma et al., 2007). Emergent properties are features of a complex system that are not present at the lower level but arise unexpectedly from interactions among the system’s components. An emergent property cannot be understood simply by examining in isolation the properties of the system’s components, but requires a consideration of the interactions among the system’s components (Kauffman, 1993; Kelso, 1995; Camazine et al., 2001; Corning, 2002). Based on this insight, group selection is the emergent behavior of complex systems. I prefer the term “community selection” instead of “group selection” because it has the connotation of “commonality”, e.g. common ecological factors (Wilson & Swenson, 2003). Communities are defined by shared interactions and common selective pressures (Ehrlich & Raven, 1964; Lubchenko & Gaines, 1981; Goodnight, 1990a, b). This is in accord with the group selection models of Wilson and Sober (Wilson 1975, 1980; Wilson & Sober 1989, 1994) in which there may be linkages (a “shared fate”) between two or more individuals (genotypes) in a randomly breeding population.

If two Newtonian forces act on a single body, say gravitation and friction, then the effects of their actions are separable. One can attribute some aspect of the final motion as due to friction, the other to gravity (Mitchell, 2009). To get the overall effect in this case the vector sum of the forces is used to predict the motion that will result from the simultaneous action of gravity and friction. Vector addition is in physics a general method for combining the effects of independent forces on the motion of a body (Mitchell, 2009). Natural selection has the attributes of a vector: force and direction (Sober, 1984). Accordingly, I envisage a selective pressure as a vector force (Sober, 1984; Eldredge, 2003) acting on an organism (but see Matthen & Ariew, 2002). Like Sober (1984), I use vectors not in the Newtonian sense but as a metaphor to illustrate evolutionary processes. Of course, the vectors acting on the individuals are not independent. In evolution, organisms that interact with each other mutually affect the strength and direction of their selection vectors and coevolve. Convergent selective pressures on individuals, visualized as a bundle of vectors that point into the same direction, should create a force field and momentum of coordinated movement. The coordinated movement of units within
communities can be found both at the cellular and behavioral level. Cells performing collective migration share many cell biological characteristics with independently migrating cells but, by affecting one another mechanically and via signaling, these cell groups are subject to additional regulation and constraints (Rørth, 2009). Thus, for collective migration, the relevant cell biology is that of a single migratory cell plus the features added by the community effects. A characteristic of collective behavior of cells found in a wound monolayer is the emergence of leaders and followers and coordination between the movement vector of one cell and its neighbors (Poujade et al., 2007; Vitorino & Meyer, 2008). Likewise, the coordinated movement of a school of fish, a raiding column of army ants, the synchronous flashing of fireflies, are emergent behaviors of complex systems (see chapter 14.3). But how could the vector forces elicit the coordinated, heritable, movement in communities of individuals? Clearly these “movers” have to be exchanged at the level of hereditary units. In fact, individuals are not that individual as is often assumed. Their individuality depends on the unique assortment of genetic modules (von Mering et al., 2003; Pereira-Leal et al., 2006). However, individuals of sexually reproducing taxa are more or less transiently assorted entities in a network of exchanged modules from a population pool. Recombination is the glue that keeps them together and that exchanges the “vectors” that drive communities into certain directions. Genetic vectors are organized in modules (Donadio et al., 1991). Intriguingly, the modularity of metabolic networks of organisms (Parter et al., 2007; Kreimer et al., 2008) and other biological systems (He et al., 2009; Lorenz et al., 2011) appears to be an evolutionary signature of variable environments. Moreover, the modular organization greatly accelerates evolution (Kashtan et al., 2007). The directional forces that are determined by ecological pressures and genomic constraints, give rise to community-level processes such as coevolution, cooperation, mutualism and symbiosis. Even convergent evolution, the concordant response of distinct communities, can be explained by the vector model.

The evolutionary reality of community-level processes that ensure the sustainability of ecosystems cannot be explained by selection at the level of selfish individuals. Broadly defined, synergy refers to the combined (cooperative) effects that are produced by two or more particles, elements, parts or organisms – effects that are not otherwise attainable (Corning, 1983, 1995, 1996, 1997, 2005). Motive forces, as visualized by vectors, drive bodies into certain directions. Natural selection is a kinetic force. Community selection can be visualized by more or less parallel vectors that act on groups of individuals representing synkinetic selection.

Both the time frame and scale of environmental fluctuations that become selectively relevant are altered in community selection vs. individual selection. Groups experience a stronger selection pressure than individuals for homeostasis with respect to reproductively limiting variables, because their greater longevity exposes them more often to suboptimal physical conditions, and greater physical size means they encompass a larger fraction of any resource/nutrient gradient. Groups achieve homeostasis by differentiation into microcosms with specialist functions, e.g. cell types. Such differentiation is more limited in individuals due to their smaller size and shorter lifespan. Hence tolerance of fluctuation in certain physical variables is proposed to be weaker in individuals than in groups (Boyle & Lenton, 2006).

15.2 Fitness as transgenerational propensity

Fitness is often estimated as r, the instantaneous rate of increase (Clutton-Brock, 1988), or R, the net reproductive rate, or simply the total number of offspring produced in an individual’s lifespan (Clutton-Brock, 1988). It is often assumed that a simple estimate of fitness is all that is needed to understand the selection pressures operating in a particular system. The organisms’ environments play a fundamental role in determining their fitness and hence the action of natural selection. 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 (Abrams, 2009a). In constant environments natural selection leads to each individual organism maximizing its expected number of descendants left far in the future. If there are no environmental fluctuations, population fitness is maximized and measured by the arithmetic mean number of surviving descendants. In evolutionary computation, the Genetic Algorithm is based on the “survival of the fittest” principle and simulates natural evolution on computer systems to solve complex problems. Individuals are selected and reproduced according to a fitness performance criterion. The fitter the individual, the higher are its chances to produce offspring. Since the process is biased towards the regions of the solution space which enclose the fittest individuals, the
evolving population gradually loses diversity and converges. After a population has converged, it is very difficult to adapt to a new optimum when the environment changes (Cobb & Grefenstette, 1993; Simões & Costa, 2002; Bui et al., 2005). Thus, premature convergence is a problem for the Genetic Algorithm as it gradually loses its exploratory ability during the evolutionary process under an oversimplified “survival of the fittest” principle. In stochastic environments (see chapter 10), the evolutionary fate of a genotype can change from generation to generation (Abrams, 2009a). The propensity definition of fitness takes the transgenerational stochasticity of fitness into account. Objective probabilistic dispositions are known as propensities (this use of the term was originated by Popper [1959]). Within the philosophy of biology, the most widely accepted modern definition of evolutionary fitness is probabilistic propensity, which holds that a trait confers fitness on an organism if that trait has the probabilistic propensity of increasing the organism’s (reproductively viable) offspring (Brandon, 1978; 1990; Mills & Beatty, 1979; Burian 1983; Richardson & Burian, 1992; Millstein, 2002; Pence & Ramsey, 2013). However, various inconsistencies and implausibilities of this concept are unresolved (Bouchard & Rosenberg, 2004; Abrams, 2009b; Ariew & Ernst, 2009).

Natural environments continuously undergo changes that alter the fitness landscapes, displacing populations towards suboptimal fitness regions. R.A. Fisher thought that environmental changes are so ubiquitous that, as he once said, Wright’s peaks and valleys are more like the undulating wave crests and troughs of an ocean than a mountainous landscape. He believed that a population rarely, if ever, finds itself in a position where no allele frequency change could increase its fitness (Crow, 1987). The static concepts of fitness and fitness landscapes (Wright, 1931, 1932; Gavrilets, 2004; Svensson & Calsbeek, 2012) have been supplemented by dynamic concepts (Wilke et al., 2004), (Mustonen & Lässig, 2009). Dubbed fitness seascapes (Mustonen & Lässig, 2010), they take the ever changing nature of environmental conditions into account. The dynamical approach leads to a quantitative measure of adaptation called fitness flux, which counts the excess of beneficial over deleterious genomic change (Mustonen & Lässig, 2009). Dobzhansky (1950), in a seminal statement on adaptation to diverse environments, wrote ‘Changeable environments put the highest premium on versatility rather than on perfection in adaptation’. Typically, individual fitness and population fitness are in conflict. While selfish behavior is favored by individual selection, cooperation is favored by population-level selection (van Baalen & Rand, 1998; Traulsen & Nowak, 2006; see Heininger, 2015). This insight is at variance with the individual-as-maximising-agent paradigm of orthodox Darwinism (Grafen, 1999). The individual-as-maximising-agent does not make sense even if looking at the level of an individual, because an individual may be displaying a behavior that is not adapted to the environment. But, it makes sense at the level of the population because the population is displaying a range of behaviors making it always adapted to the environment. Therefore, while the individual is not the most fit, the population is (Dubravcic, 2013). This has been shown in bacteria that change between fast growing/antibiotic sensitive and slow growing/antibiotic resistance states (Balaban et al., 2004), B. subtilis expressing sporulating and non-sporulating state (Veening et al., 2008a, b), plant seeds that germinate at different time points (Simons, 2009), etc. (see chapter 11).

In a constantly changing and resource-limited environment, fitness is defined by reproduction rather than survival of the individual. In fact, survival is only evolutionarily relevant in the tautological sense of “survive to reproduce”. Lonesome George, “the rarest living creature” according to the Guinness Book of World Records, the apparent sole survivor of the now probably extinct Geochelone abingdoni species of giant Galápagos tortoises from Pinta Island left no offspring and, although surviving for approx. 100 years, had an overall Darwinian fitness of zero. Therefore, I advocate to delete the term survival altogether from fitness definitions. A trait that enhances an organism’s viability, but renders it sterile, has an overall fitness of zero. This includes transgenerational processes such as the mutation “grandchildless” in Drosophila (Boswell et al., 1991) and C. elegans mutations that end in sterility not one or two, but dozens of generations later (Ahmed & Hodgkin, 2000). On the other hand, a trait that slightly reduces viability while augmenting fertility, may be very fit overall (Sober, 2001). There are short-term and long-term aspects to fitness (Beatty & Finsen, 1989; Sober, 2001; Pence & Ramsey, 2013). This distinction is not trivial. In fact, short-term reproductive success may threaten the evolutionary success of a geno-/phenotype, by placing too great a demand on available resources (Beatty & Finsen, 1989). Accordingly, a prudent resource management is routinely observed in wild populations (see chapter 15.3.1). Clearly, a population is doomed that although able to reproduce a million times is unable to sufficiently protect its offspring against e.g.,
predators, until reproductive maturity. In many taxa, the survival of offspring is dependent on parental care, which is defined as any trait that enhances the fitness of offspring and originated/is maintained for this function (Smiseth et al., 2012). Parental care is common across animal taxa and increases offspring survival and/or quality in a range of species (Clutton-Brock, 1991; Smiseth et al., 2012).

Current concepts of fitness put much emphasis on the representation of genes in the next generation. Taking into account that evolution is an iterative process, long-term concepts of fitness (Thoday, 1953; Cooper, 1984; Beatty & Finsen, 1989; Sober, 2001; McNamara et al., 2011; Pence & Ramsey, 2013) suggest that fitness should be defined as the probability of leaving descendants in the long run. Asymmetric fitness curves combined with temporal environmental fluctuations can lead to strategies that appear to be suboptimal in the short-term, but are in fact optimal in the long run (Ruel & Ayres, 1999; Martin & Huey, 2008). The reason for the apparent departure from optimality is that deviations to the right of the fitness peak reduce fitness more than equivalent deviations to the left do. Gillespie (1973a, b), Hartl and Cook (1973) and Karlin and Liberman (1974, 1975) first showed that the evolution of a system under temporal fluctuations is determined not only by expected fitness in a given generation, but also by the degree of variation in fitness over time, and established the geometric mean fitness principle (Lande, 2008; Frank, 2011). It states that in a random environment, alleles that increase the geometric mean fitness can invade a randomly mating population at equilibrium.

The obvious reason to be suspicious of the idea that variability has been fine-tuned in order to maximize the evolutionary potential of populations is that it suggests a teleological view of evolution. Natural selection cannot adapt a population for future contingencies any more than an effect can precede its cause, so any future utility of the capacity to generate variation can have no influence on the maintenance of that capacity in the present. As Sydney Brenner supposedly remarked many years ago, it would make no sense for a population in an early geological period to retain a feature that was useless merely because it might “come in handy in the Cretaceous!” Teleology need not be invoked to support evolvability arguments, however. A history of environmental uncertainty could favor a population with increased variability over others because such a population is more successful at adapting. Sniegowski and Murphy (2006) called this the evolvability-as-adaptation hypothesis. In fact, there is a great amount of evidence suggesting that evolvability itself is a selectable trait and hence, evolvability evolves (Wagner & Altenberg, 1996; Turney, 1999; Partridge & Barton, 2000; Bedau & Packard, 2003; Woods et al., 2003; Earl & Deem, 2004; Jones et al., 2007; Colegrave & Collins, 2008; Crombach & Hogeweg, 2008; Draghi & Wagner, 2008; Pigliucci, 2008; Palmer & Feldman, 2011; Pavlicev et al., 2011; Woods et al., 2011). Both temporal and spatial environmental variation can select for evolvability (van Nimwegen et al., 1999; Wilke et al., 2001b; Siegal & Bergman, 2002; de Visser et al., 2003; Wagner, 2008; Palmer & Feldman, 2011) and can speed up evolution (Kashtan et al., 2007; Parter et al., 2008; Draghi & Wagner, 2009).

Increasing evolvability implies an accelerating evolutionary pace (Turney, 1999). For evolvability to increase, environmental change must occur within certain bounds. If there is too little change, there is no advantage to evolvability. If there is too much change, evolution cannot move fast enough to track the changes (Turney, 1999). RNA virus genotypes with similar fitness may differ in their evolvability (Burch & Chao, 2000; McBride et al., 2008). To understand what determines the long-term fate of different clones, each carrying a different set of beneficial mutations, Woods and co-workers (2011) “replayed” evolution by reviving an archived population of Escherichia coli from a long-term evolution experiment and compared the fitness and ultimate fates of four genetically distinct clones. The expected scenario was that eventual winners (EW) clones were already more fit than eventual losers (EL) clones at generation 500, but competition experiments showed that actually the opposite was the case. Surprisingly, two clones with beneficial mutations that would eventually take over the population after 1,500 generations had significantly lower competitive fitness after 500 generations than two clones with mutations that later went extinct. Replaying the experiment many times starting with the 500-generation EWs and ELs showed that the EWs indeed beat the ELs most of the time. Likewise, E. coli strains with larger fitness defects due to deleterious mutations are more evolvable than wild-type clones in terms of both the beneficial mutations accessible in their immediate mutational neighborhoods and integrated over evolutionary paths that traverse multiple beneficial mutations (Barrick et al., 2010).

15.3 Reproductive fitness in stochastic environments

The assumption that expected or within-generation fitness is maximized by natural selection is simply wrong. Simons, 2002
In stable environments (see chapter 9), the default setting of orthodox Darwinism, short-term fitness predicts long-term fitness. Hence current concepts of fitness put much emphasis on the individual's representation of genes in the next generation. However, theory predicts that the fitness of a life-history strategy may be considerably different in a random environment compared with a constant environment or in populations with and without density dependence (Tuljapurkar 1989, 1990a, b; Mueller et al., 1991; Kawecki, 1993; Mylius & Diekmann, 1995). Therefore, finding that a particular life-history strategy is maladaptive may be the result of oversimplistic assumptions about the ecology of the study population. In particular, when there is density-dependent regulation in a population, the fitness of one life history may depend on other life histories present in the population. In stochastic environments, the variance of selection, or more generally the entire probability distribution of fitness, becomes a critical factor of selection (Yoshimura & Shields, 1987). Dempster (1955) introduced the model in which temporal fluctuations in reproductive success for competing genotypes favor the genotype with the highest geometric-mean reproductive success. Ever since, the standard criterion for evaluating the fitness of genotypes in stochastic environments is the geometric mean of the growth rates (geometric mean fitness) (Dempster, 1955; Cohen, 1966; Lewontin & Cohen, 1969; Kuno, 1981; Klinkhamer et al., 1983; Metz et al., 1983; Frank & Slatkin, 1990; Yoshimura & Clark, 1991; Yoshimura & Jansen, 1996; Hopper, 1999; Simons, 2009; Yoshimura et al., 2009). Geometric mean fitness is a concept widely used in ecology and evolutionary biology to understand persistence of populations in fluctuating environments (Lewontin & Cohen, 1969; Levins, 1969; Gillespie, 1974a, b; Kuno, 1981; Yoshimura & Jansen, 1996; Jansen & Yoshimura, 1998).

Risk-sensitive reproductive strategies may reduce the average (arithmetic mean) of individual reproductive output, while yet maximizing the population geometric mean; this trade-off in terms of average reproduction is 'bet hedging' (Gillespie, 1973, 1974a; Slatkin, 1974; Seger & Brockmann, 1987; Philippi & Seger, 1989). Risk-sensitive behavior is variance-sensitive behavior (Smallwood, 1996; Ydenberg, 2007; Mayack & Naug, 2011; Ratifkainen, 2012). However, the notion of a 'sacrifice' of expected fitness for geometric mean fitness is deceptive. There is no detrimental effect of maximizing the geometric mean fitness and, hence, no trade-off between the mean and variance in fitness exists; the assumption that expected or within-generation fitness is maximized by natural selection is simply wrong (Simons, 2002). There is typically a trade-off between the number of surviving offspring and their quality (i.e. their reproductive value) so that in general fitness is not maximized by maximizing the mean number of surviving offspring (McNamara et al., 2011; Heininger, 2013). A rigorous definition of bet-hedging includes lower expected arithmetic mean fitness, as well as greater expected geometric mean fitness (Seger & Brockmann, 1987; Simons, 2002, 2009). Bet-hedging involves a trade-off between the mean and variance of fitness. If the environment varies temporally, phenotypes with low variances of fitness may be favored over alternatives with higher variances and higher mean fitnesses (Philippi & Seger, 1989). This reduction in among-generation variation in fitness (yielding a higher geometric mean) forms the basis of bet-hedging theory: bet-hedgers, reducing variance in fitness, don't necessarily do best all the time, but they perform most consistently and are therefore favored by selection (Cohen, 1966; Roff, 1992). Thus, in stochastic environments, individual fitness maximization regimes are replaced by population-level fitness maximization strategies that yield suboptimal fitness results for individuals (Cohen, 1966; Ellner, 1986; McNamara 1995; 1998; McNamara et al., 1995; Yoshimura & Jansen, 1996). Geometric mean fitness is a typical example of such a selection criterion under environmental stochasticity over many generations (Lewontin & Cohen, 1969; Yoshimura & Clark, 1991). The total resources available to the population limit reproductive success. Density-dependent competition causes the reproductive success of each type to be influenced by the reproduction of other types. For that reason, one cannot simply multiply the reproductive successes of each type independently and then compare the long-term geometric means. Instead, each bout of density-dependent competition causes interactions between the competing types. Those interactions depend on frequency (Frank, 2011).

When the fitness of a genotype varies over generations, the appropriate measure of its relative growth rate is its geometric mean fitness, rather than its arithmetic mean fitness. Lewontin and Cohen (1969) presented a formal argument showing the absurdity of the use of the arithmetic mean fitness under environmental variability: even when expected fitness approaches infinity, the probability of extinction in a variable environment may rise to one. Fitness, like return on investment, is determined by a multiplicative process (Dempster, 1955; Gillespie, 1974a)—that of reproduction—and bet hedging increases the geometric-mean fitness (the nth root of the product of...
n fitness values) by reducing fitness variance over generations (Gillespie, 1977). If fitness for a given genotype is zero in generation z (i.e. goes extinct in that generation), then the fitness of that genotype across generations x, y, z is not the arithmetic mean of the fitness of these three generations, but zero. If the numbers vary, then the geometric mean is always less than the arithmetic mean; in general, the geometric mean becomes smaller as the numbers being averaged become more variable. Thus the geometric mean fitness of a genotype can be increased by reducing the variance of its fitness (over generations), even if the reduction of variance also entails a reduction of the arithmetic mean. When fitness fluctuates through time and the fluctuations are modest, the identity of the allele that predominates in a population depends on both the mean and the variance in fitness. Consequently, if two alleles have the same (arithmetic) mean fitness through time, the allele that ‘wins’ is the one with the smaller variance in fitness. Thus, it is advantageous for alleles to avoid large fluctuations in fitness. When there are no fluctuations in fitness through time (constant environments), the geometric mean fitness collapses to the arithmetic mean fitness (Orr, 2009).

Jensen’s inequality (1906), a mathematical property of nonlinear functions (Ruel & Ayres, 1999), provides a fundamental tool for understanding and predicting consequences of variance, but it is only just beginning to be explicitly acknowledged in the primary literature (Stockhoff, 1993; Smallwood, 1996; Anderson et al., 1997; Karban et al., 1997; Ruel & Ayres, 1999; Martin & Huey, 2008; Lof et al., 2012). Asymmetric fitness curves are probably common given that many ecological and physiological processes affecting fitness are likely to exhibit skewness, particularly with respect to temperature (Gilchrist, 1995; Martin & Huey, 2008; Dell et al., 2011). Asymmetric fitness curves combined with temporal environmental fluctuations can lead to strategies that appear to be suboptimal in the short-term, but are in fact optimal in the long run (Ruel & Ayres, 1999; Martin & Huey, 2008). The reason for the apparent departure from optimality is that deviations to the right of the fitness peak reduce fitness more than equivalent deviations to the left do.

15.3.1 Reproductive prudence

The tragedy of the commons (a situation where individual competition reduces the resource over which individuals compete, resulting in lower overall fitness for all members of a group or population) provides a useful analogy allowing to understand why shared resources tend to become overexploited (Hardin, 1968). The logic of the tragedy of the commons predicts that individual good will be maximized with disastrous consequences for the population. Overexploitation of resources can result in reduced per capita birth rates or increased mortality and thereby provides an upper limit to population size (Hairston et al., 1960; Arcese & Smith, 1988). If there are time-lags involved, this mechanism might also result in periodic oscillations around a ‘carrying capacity’ (McCaulay et al., 1999). The tragedy of the commons analogy has become increasingly used to explain why, in principle, selfish individuals in a multitude of parasite, animal and plant populations evolved means to avoid the overexploitation of limited collective resources (Frank, 1995; Gersani et al., 2001; Falster & Westoby, 2003; Foster, 2004; Wenseleers & Ratnieks, 2004; Rankin & López-Sepulcre, 2005; Kerr et al., 2006; Rankin & Kokko, 2006; Mideo & Day, 2008; Carter et al., 2014). Factors such as high relatedness in social groups (Wenseleers & Ratnieks, 2004), diminishing returns (Foster, 2004), policing and repression of competition (Frank, 1995, 1996a; Hartmann et al., 2003; Ratnieks & Wenseleers, 2005; Kentzoglukanakis et al., 2013), altruism (Frank, 1996b; van Baalen, 2002; Lion & van Baalen, 2008), reputation (Milinski et al., 2002), pleiotropy (Foster et al., 2004), plasticity (Fischbacher et al., 2012; Cavaliere & Poyatos, 2013) or control of population density (Hauert et al., 2006; Kokko & Rankin, 2006; Rankin, 2007; Frank, 2010) have been argued to constrain the evolution of overexploitative behavior, and thus reduce the potential for a tragedy of the commons to arise in such populations.

15.3.1.1 Geometric mean fitness criterion

In unstable environments, the geometric mean is always lower than the arithmetic mean (see also chapter 15.3). In fluctuating environments, when geometric mean fitness is maximized, individual optimization fails (Cohen, 1966; Ellner, 1986; McNamara 1995; 1998; McNamara et al., 1995; Yoshimura & Jansen, 1996). Under the geometric mean criterion, behavior appears to be determined largely by a worst case scenario; behavior may appear suboptimal under the perspective of normal or average conditions (Yoshimura & Clark, 1991; Yoshimura & Jansen, 1996). In other words, in unpredictable environments it is better to do on average bad but stable as opposed to sometimes good and sometimes bad (Starrfelt, 2011). Except under extreme environmental conditions, mammalian litters (Murie & Dobson, 1987; Risch et al., 1995) and avian clutches (Perrins, 1965; Klomp, 1970; Murray, 1979; Lessells, 1986; Murphy & Haukojoa, 1986; Boyce & Perrins, 1987; Vander Werf, 1992) larger than those that are
observed in nature might result in increased fecundity, with little if any cost of reproduction in terms of parental survival. However, in unusually bad years such large clutches might be disastrous, in terms of parental survival (Yoshimura & Clark, 1991; Yoshimura & Shields, 1992). According to David Lack’s (1947, 1954) brood reduction hypothesis, asynchronous hatching facilitates adaptive brood reduction when environmental conditions are poor, and thus maximizes the number of fledglings produced under such circumstances (Forbes, 1991; Amundsen & Slagsvold, 1998). An illustrative example was given by Philipp & Seger (1989): “Suppose that years are ‘good’ or ‘bad’ with equal probability, and that the wild type produces, on average, 9 offspring in good years and 1 offspring in bad years, for an average of 5. Now introduce a mutant that produces 5 offspring in good years and 3 offspring in bad years, for an average of only 4. Despite its lower mean fitness, the mutant quickly goes to fixation because its geometric mean fitness (3.87) is much higher than that of the wild type (3.0) and its variance lower. The mutant’s best performance is much worse than the wild type’s best, but its worst is better, and this is the key to its success.” Evidence for this prudent reproduction is found in all taxa.

15.3.1.2 Viruses

In a study, groups of bacteria and bacteria-infecting viruses were grown in 96 separate wells on plates. “Migration” between the groups was executed by a robot transferring small quantities of liquid between wells according to prespecified schemes. Under biologically plausible migration schemes, “prudent” virus strains were able to outcompete more “rapacious” strains, despite their selective disadvantage within each group. Prudent phage dominate when migration is spatially restricted, while rapacious phage evolve under unrestricted migration (Kerr et al., 2006).

15.3.1.3 Microbes

One characteristic of bacteria is that microbial growth yields are often 50% less than the optimal yield (Westerhoff et al., 1983). There is an inevitable thermodynamic trade-off between growth rate and yield among heterotrophic organisms (Pfeiffer et al., 2001; Novak et al., 2006). Two opposing ecological strategies exist at either end of the growth rate/yield spectrum: a fast-growing, low yield competitive strategy and a slow-growing, high yield cooperative strategy (Pfeiffer et al., 2001; Kreft & Bonhoeffer, 2005). Metabolic pathways are faced with a trade-off between the rate and yield of ATP production. Simple evolutionary models argue that this trade-off generates a fundamental social conflict in microbial populations: average fitness in a population is highest if all individuals exploit common resources efficiently, but individual reproductive rate is maximized by consuming common resources at the highest possible rate (MacLean & Gudelj, 2006; MacLean, 2008). For microbes, the cooperative, slow, efficient growth strategy is more successful in spatially structured environments such as biofilms (Pfeiffer et al., 2001; Kreft, 2004; Kreft & Bonhoeffer, 2005; MacLean & Gudelj, 2006).

Gene expression noise is a selected-for trait, particularly to increase survival in stressful conditions (see chapters 11.1.1 and 11.3). Within the conceptual framework of traditional Darwinism it is hard to understand that gene expression noise in yeast reduces the mean fitness of a cell by at least 25%, and this reduction cannot be substantially alleviated by gene overexpression (Wang & Zhang, 2011). However, within the framework of the cybernetic model of evolution this trade-off between growth in benign conditions and survival in stressful conditions makes perfect sense. This trade-off illustrates that the geometric mean fitness criterion can also be applied to microbes (Beaumont et al., 2009; Ratcliff & Denison, 2010).

From an evolutionary perspective, mechanistic coupling between transmission and virulence strongly shapes the life history of parasites (Day, 2002; Frank & Schmid-Hempel, 2008). A fundamental property underlying many perspectives on the evolution of virulence is a link or ‘trade-off’ between the virulence of an infection and the reproductive capacity of the parasite (Anderson & May, 1982; May & Anderson, 1983; Ewald, 1987; Bull, 1994; Frank, 1996). The most commonly assumed mechanism for this trade-off is that virulence is an unavoidable consequence of parasite reproduction in the host and hence that higher parasite reproduction results in higher virulence. A parasite’s fitness improves with increases in its reproductive capacity, but is diminished by high virulence because virulence debilitates the host’s ability to transmit the parasite (Messinger et al., 1999). Highest parasite fitness is thus achieved as a compromise, the exact optimum depending on the shape of the trade-off surface. Comparisons across parasites evolved in nature and from selection experiments are consistent with trade-offs (Bull et al., 1991; Diffley et al., 1987; Darsley et al., 1990; Day et al., 1993; Herre, 1993; Ebert, 1994; Ewald, 1994; Ebert & Mangin, 1997; Turner et al., 1998; Mackinnon & Read, 1999; Messenger et al., 1999; Paul et al.,
2004; Salvadoron et al., 2005; de Roode et al., 2008; Mackinnon et al., 2008). Many processes such as pathogen adaptation to within-host competition, interactions with the immune system and shifting transmission routes, will all be interrelated to virulence-transmission trade-off making sweeping evolutionary predictions harder to obtain (Alizon et al., 2009). If host immunity is short-lived, or if it is imperfect, the level of host exploitation should increase. Only parasites causing diseases with long-lived immunity are likely to be prudent in space (Lion & Boots, 2010).

Testing Cohen's (1966) classic bet-hedging model using the fungus Neurospora crassa, Graham et al. (2014) allowed ascospore dormancy fraction in N. crassa to evolve under five experimental selection regimes that differed in the frequency of unpredictable 'bad years'. The straightforward prediction of the model is that, by the geometric-mean principle, dormancy fraction should evolve to equal the probability of occurrence of a bad year. By contrast, the prediction of the arithmetic-mean principle is the evolution of zero dormancy (immediate germination) under a broad range of ecological scenarios; namely if the probability of a good year is greater than 0.5 (Graham et al., 2014). Results were consistent with bet-hedging theory: final dormancy fraction in 12 genetic lineages across 88 independently evolving samples was proportional to the frequency of bad years, and evolved both upwards and downwards as predicted from a range of starting dormancy fractions (Graham et al., 2014).

15.3.1.4 Prolonged dormancy

In many insect species (Waldbauer, 1978; Ushatinskaya, 1984; Tauber et al., 1986; Danks, 1987, 1992; Hanski, 1988; Menu, 1993a, b; Menu & Debouzie, 1993; Roux et al., 1997; Danforth, 1999; Menu et al., 2000) but also in other organisms such as plants (e.g. Venable & Lawlor, 1980; Venable, 1989; Philipp, 1993a, 1993b; Clauss & Venable, 2000), crustaceans (Ellner & Hairston, 1994; Hairston et al., 1995, 1996b) and tropical fishes (Wourms, 1972), life cycle duration varies within the population. Certain individuals of the same generation reproduce after 1 year and others after 2 or more years because of prolonged dormancy. Diapause is a genetically programmed developmental response that occurs at a specific stage for each species that allows synchronization of the life cycle with seasonal variations in the environment (Tauber et al., 1986; Danks, 1987, 1992). Interestingly, diapause lasting more than 1 year, namely “prolonged” or “extended” diapause, is not exceptional for insect species (Danks, 1987, 1992; Hanski, 1988). It usually occurs in populations whose seasonal resources fluctuate unpredictably in abundance and availability (Hanski, 1988). Species undergoing prolonged dormancy (diapause or quiescence) usually reside in arid or semiarid areas (Nakamura & Ae, 1977; Sims, 1983; Powell, 1987, 1989, 2001; Danforth, 1999; Tauber & Tauber, 2002) as well as in regions of the arctic zone (Danks, 2004). Nonetheless, prolonged dormancy has also been reported in temperate zone species (Barnes, 1952; Neilson, 1962; Prentiss, 1976; Shapiro, 1979, 1980; Annila, 1982; Hedlin et al., 1982; Tzanakanis et al., 1991; Levine et al., 1992; Menu, 1993a, b; Menu & Debouzie, 1993; Higaki & Ando, 1999; Maeto & Ozaki, 2003; Higaki, 2005; Matsuo, 2006; Wang et al., 2006; Chirumamilla et al., 2008). Prolonged diapause is a within-generation bet-hedging phenotype (Menu & Desouhant, 2002). As noted by Hutchinson (1996), “Biologists who are used to thinking in terms of maximisation of individual fitness are often perturbed that a seed (or an insect) should agree not to germinate (emerge as adult) immediately when its own chances of reproducing are lower if it spends a year in dormancy.” Individuals that express prolonged dormancy are exposed to increased mortality risks and they postpone reproduction, both of which may result in fitness costs (Danks, 1987; Leather et al., 1993; Hairston, 1998). In an experimental study, prolonged dormancy did not affect adult longevity but both lifetime fecundity and oviposition were significantly decreased (Moraiti et al., 2012).

15.3.1.5 Social insects

Several studies reported a survival advantage of multiple foundress colonies compared with single foundress colonies of the wasp genera Polistes (Metcalf & Whitt, 1977; Gibo, 1978; Tibbetts & Reeve, 2003), Belonogaster (Keeping & Crewe, 1987; Tindo et al., 1997a), and Ropalidia (Shakarad & Gadagkar, 1995). Alldapine bees (Hogendoorn & Zamm, 2001) and social shrimps (Duffy, 2002). In the primitively eusocial wasp species Belonogaster juncea juncea multiple foundress colonies were significantly more successful than single foundress colonies in producing at least one adult (Tindo et al., 2008). The total productivity of the colonies increased significantly with the number of associated foundresses, but the productivity per capita did not. No single foundress colony (out of 13) reached the sexual phase, while eight (out of 36; 21.6%) multiple foundress colonies did. The increase in total productivity as a function of group size is in line with previous findings reported on primitively eusocial species (Michener, 1964;
Shakarad & Gadagkar, 1995; Tindo et al., 1997b; Tibbetts & Reeve, 2003). On the other hand, the decreasing per capita productivity concomitant with an increasing number of females noted in the study of Tindo et al. (2008) illustrates Michener’s paradox (1964) in primitively eusocial insects (Michener, 1964; Noonan, 1981; Strassmann et al., 1988; Shakarad & Gadagkar, 1995; Gadagkar, 1996; Hogendoorn & Zammit, 2001; Seppä et al., 2002; Soucy et al., 2003). The coefficient of variance of the per capita productivity significantly decreased with group size, as Wenzel and Pickering (1991) noted in the model they created to explain the paradox (Tindo et al., 2008). Wenzel and Pickering (1991) suggested that individuals in larger groups might trade lower per capita productivity for less variability and greater predictability.

15.3.1.6 Prudent predators

Predator-prey systems (and related host-pathogen systems) have been studied theoretically for decades. Most of the studies have focused on the role of environmental stochasticity, the relevance of nonlinear interactions or of spatial effects, to explain the mechanism of cycling (Nisbet & Gurney, 1982; Renshaw, 1991; Kaitala et al., 1996; Aparicio & Solari, 2001; Bjørnstad & Grenfell, 2001; Pascual et al., 2001; Pascual & Mazzeega, 2003). The “prudent predator” concept (Slobodkin, 1961, 1974; Goodnight et al., 2008) has elucidated evolutionary outcomes of predator-prey interactions and provided evolutionary mechanisms to resolve the tragedy of the commons dilemma. A predator here is defined as any species that consumes or exploits another species in order to survive and reproduce, including pathogens, parasites, parasitoids, grazers and browsers, as well as “true” predators. The classical model of predator-prey dynamics, the Lotka-Volterra equation, predicts that under most conditions predator populations, like prey populations, due to overexploitation of resources, go through a series of oscillations between feast and famine, at each cycle approaching the brink of extinction (Holland, 1995; Mitteldorf, 2010). The paradox is that in the natural world, we know that predator and prey stably coexist in nature even when heritable variation exists for traits involved in predator attack rates (e.g. Forsman & Lindell, 1993; Virol et al., 2003; Palkovacs & Post, 2008). The overexploitation of resources can only be prevented by conservation of the resource by prudent reproduction (Slobodkin, 1961, 1974; Goodnight et al., 2008). Evidence for the evolutionary merit of reproductive prudence comes from multiple experimental studies in various taxa that is supported by theoretical models (Gilpin, 1975; Nathanson, 1975; Wilson, 1978; Wade, 1980a; Holmes, 1983; Walker, 1984; Rand et al., 1995; Savill & Hogeweg, 1998; Sober & Wilson, 1998; Boots & Sasaki, 2000; Haraguchi & Sasaki, 2000; Rauch et al., 2002, 2003; Werfel & Bar-Yam, 2004; Kerr et al., 2006; Goodnight et al., 2008; MacLean, 2008; Borrello, 2012; Carter et al., 2014). Reproductive prudence of cells arose as a necessary prerequisite of multicellularity (Buss, 1987; Maynard Smith & Szathmáry, 1995; Frank & Nowak, 2004). Cancerogenesis can be regarded as a violation of this reproductive prudence resulting in the tragedy of the commons (Nunney, 1999; Stoler et al., 1999).

...that an opinion has been widely held is no evidence whatever that it is not utterly absurd….

Bertrand Russell (1929)

16. Life history phenotypes of bet-hedging

16.1 Turnover of generations: bet-hedging in time?

Turnover of generations may be considered as transgenerational bet-hedging. Traditional concepts of aging (the “evolutionary theories of aging”) do not take into account the stochasticity of environments. In fluctuating environments it cannot be expected that fitness of individuals is optimal over longer intervals. Theoretical studies on variability in life cycle duration both in plants (Philippi, 1993a, 1993b; Clauss & Venable, 2000) and insects (Danforth, 1999; Menu et al., 2000) proposed bet-hedging as an explanation of such variability. As adaptation to environmental unpredictability, diapause cycle length must be expressed as a responsiveness to unpredictable proximate environmental factors (i.e. factors without predictive value for the decision at hand) (Menu 1993; Menu & Debouzie, 1993; Menu et al., 2000; Menu & Desouhant, 2002).

Life history traits of long-lived vertebrates constrain the ability of populations to respond to environmental perturbations resulting in chronic increases in mortality (Heppell et al., 2000; Fordham et al., 2007) because compensatory responses are thought to be limited and recovery is slow (Musick et al., 2000). A ‘slow–fast’ continuum in life histories exists for a range of taxa (Heininger, 2012), including mammals (Heppell et al., 2000), birds (Sæther et al., 1996), reptiles (Webb et al., 2002) and sharks (Smith et al., 1998), and a species’ position along this continuum influences how populations will respond to change in a demographic trait (Sæther & Bakke, 2000). Intriguingly, organisms that live either in more stable environments such as...
the deep sea or are more resilient to environmental perturbations and hence can tolerate a broader range of environmental conditions (such as endothermic organisms) have greater longevities (Finch, 1990).

16.2 Iteroparity

To understand how environmental fluctuations shape the evolution of life histories, stochastic demography has to be used (Tuljapurkar, 1990b; Caswell, 2001; Tuljapurkar et al., 2009). Distributing reproduction in time has been visited by many studies since Cole (1954) coined the terms “semelparity” and “iteroparity”. Cole (1954) viewed iteroparity as a paradox because semelparity, a single bout of reproduction, should always be favored in a constant environment by the compounding nature of exponential growth. In a constant environment, Cole’s paradox boils down to a question of lifetime reproductive success; the type producing most offspring over a lifetime will come to dominate (Mylius & Diekmann, 1995; Zeineddine & Jansen, 2009). As discussed in chapter 10, reproductive success can be highly variable (Hastion et al., 1996a). Cole’s paradox has been resolved by numerous models demonstrating that variation in reproductive success favors iteroparity (Murphy, 1968; Gadgil & Bossert, 1970; Saucier, 1974; Bulmer et al., 1974; Bell, 1976, 1980; Goodman, 1984; Bulmer, 1985; Orzack, 1985, 1993; Brashaw, 1986; Roedtink, 1987; Orzack & Tuljapurkar, 1989, 2001; Fox, 1993; Charlesworth, 1994; Cooch & Ricklefs, 1994; Erikstad et al., 1998; Benton & Grant, 1999; Brommer et al., 2000; Ranta et al., 2000a, 2000b, 2002; Katsukawa et al., 2002; Wilbur & Rudolf, 2006). Thus, life history theory holds that in the face of annual resource variability, organisms should shift from semelparous to iteroparous reproductive patterns (Murphy, 1968; Bulmer 1985, Orzack & Tuljapurkar, 1989); and furthermore, under certain circumstances they should evolve a longer lifespan and reduced annual reproduction (Stearns, 1976; Gillespie, 1977; Roff, 2002; Nevoix et al., 2010). By this theory, bet-hedging evolves to reduce the probability of investing too much in reproduction during resource-poor years, which may ultimately result in null fitness. However, for logistical reasons, the theoretical prediction that environmental variability will lead to the evolution of longer life span (Murphy, 1968; Roff, 2002) has rarely been tested or detected in wild populations (Roff, 2002; Nevoix et al., 2010).

16.3 Polyandry

Whereas for males reproductive success is expected to increase linearly with the number of mates, the advantages of multiple mating for females are less clear (Yasui, 1997; Jennions & Petrie, 2000). Mating can be costly to females in terms of time and energy, or because of increased risk of predation, injury or infection (Daly, 1978; Chapman et al., 1995; Blanckenhorn et al., 2002). Polyandry (multiple female mating) is common in a wide variety of animal taxa (Birkhead, 2000; Jennions & Petrie, 2000; Uller & Olsson, 2008). The evolutionary rationale for this behavior may differ between species and a multitude of mutually non-exclusive theories have been forwarded to explain its occurrence. For instance, polyandry may represent the combined effect of mate-encounter frequency and conflict over mating rates between males and females driven by large male benefits and relatively small female costs resulting in “convenience polyandry” (DiBattista et al., 2008; Uller & Olsson, 2008). On the other hand, polyandry may be another within-generation bet-hedging behavior (Yasui, 1998; Hopper et al., 2003; Sarhan & Kokko, 2007). There are two ways in which polyandry could be favored by bet-hedging (Jennions & Petrie, 2000). First, females may only be able to distinguish broad categories of males due to perceptual errors in assessment; or there may only be a few discrete levels of signaling by males, despite continuous variation in male quality (Johnstone, 1994). Second, there may be temporal fluctuations in the environment that lead to variable selection on fitness-enhancing traits under natural selection (e.g. Jia & Greenfield, 1997). As such, females cannot identify the male with the best viability genes for the future. In both cases, females can reduce the variance in mate quality by mating with several males whom they perceive to be broadly genetically suitable as mates (so-called ‘genetic bet-hedging’; Watson, 1991). Genetic bet-hedging (Gillespie 1973a, 1974a, 1975, 1977; Seger & Brockman, 1987; Hopper, 1999) could explain polyandry, especially when females mate indiscriminately (Yasui, 1998, 2001; Fox & Rauter, 2003). Another advantage may be a form of diversified bet-hedging akin to not putting all your eggs in one basket (Kaplan & Cooper, 1984). Benefits of polyandry may include genetic bet-hedging against environmental uncertainty, mating with costly males and genetic incompatibility (Loman et al., 1988; Watson, 1991; Zeh & Zeh, 1996; Newcomer et al., 1999; Jennions & Petrie, 2000; Yasui, 2001; Fox & Rauter, 2003; Lorch & Chao, 2003; Mäkinen et al., 2007; Byrne & Roberts, 2012). Field studies of vertebrates suggest, and laboratory experiments on invertebrates confirm, that even when males provide no material benefits, polyandry can enhance offspring survival and fitness (Madsen et al., 1992, 2005; Tregenza & Wedell, 1998, 2000, 2002; Jennions & Petrie, 2000; Zeh & Zeh, 2001, 2006; Garant et al.,
2005; Ivy & Sakaluk, 2005; Fisher et al., 2006; Sarhan & Kokko, 2007; Byrne & Whiling, 2011; Aguirre & Marshall, 2012; Reding, 2014; Garcia-Gonzalez et al., 2015). Genetic variation among sexually produced siblings could also reduce the likelihood of disease and predation (Wolfe, 1985; Zhu et al., 2000; Jokela et al., 2009; Aguirre & Marshall, 2012).

Female birds may benefit indirectly from extra-pair mating by enhancing the genetic quality of their offspring, through good genes or genetic compatibility effects (Jennions & Petrie, 2000; Kokko, 2001; Griffith et al., 2002; Neff & Pitcher, 2005). Supporting this idea, recent studies have identified a range of fitness-related traits for which extra-pair offspring are superior to their within-pair half-siblings (Hasselquist et al., 1996; Kempenaers et al., 1997; Sheldon et al., 1997; Johnsen et al., 2000; Charmantier et al., 2004; Schmoll et al., 2005; Freeman-Gallant et al., 2006; Garvin et al., 2006; Bouwman et al., 2007; O’Brien & Dawson, 2007; Dreiss et al., 2008; Fossay et al., 2008; Losdat et al., 2011).

A recent study (Gohli et al., 2013) found that more promiscuous species of passerine birds had higher nucleotide diversity at autosomal introns, but not at Z-chromosome introns. In more promiscuous species, major histocompatibility complex class IIB alleles had higher sequence diversity, and therefore should recognize a broader spectrum of pathogens. The results suggest that female promiscuity in passerine birds targets a multitude of autosomal genes for their nonadditive, compatibility benefits. Also, as immunity genes seem to be of particular importance, interspecific variation in female promiscuity among passerine birds may have arisen in response to the strength of pathogen-mediated selection (Gohli et al., 2013).

16.4 Sexual reproduction

The immune system maintains a living system’s organization by destroying parasitic bodies, such as bacteria or cancer cells. It achieves this by producing antibodies that attach themselves to the alien bodies and thus neutralize them. To find the right type of antibodies, the immune system simply produces an astronomical variety of different antibody shapes. However, only the ones that “fit” the invaders are selected and reproduced in large quantities (Heylighen, 2001). Thus, Abel (2012) thought that the only system that seems to waste energy deliberately exploring randomness is the immune system. To prepare for exposure to an indefinite array of possible antigens, the immune system must be prepared to deal with any possible new combination of viral, bacterial, mycotic, or other parasitic invasion. The immune system is unique in its continuing perusal of potential genetic sequence space and three dimensional phase space (Abel, 2012). Likewise, however, in a stochastic environment the best strategy to increase fitness is to take every possible path at every next step. As a result, no configurations should be missed (Fu, 2007). Thus environmental stochasticity elicits bet-hedging as risk-spreading response resulting in (epi)genetic, developmental, phenotypic, physiological and behavioral variation on which selection can act (figure 1C).

Several theoretical models indicate that sexual reproduction is selected for in variable environments (Hines & Moore, 1981; Weinshall, 1986; Roughgarden, 1991; Robson et al., 1999). Sexual reproduction is the ultimate bet-hedging enterprise and its evolutionary success the selective signature of stochastic environments (Heininger, 2013). Sexual reproduction subjects an extremely large variety of germline cells that are organized like a quasispecies to a cascade of selective regimes before the most resilient (Holling, 1996) are released and exposed to natural selection (Heininger, 2013). With these features, gametogenesis in sexually reproducing organisms is characterized as complex self-organized system as described by Heylighen (2001): “The system needs a fitness criterion for choosing the best action for the given circumstances. The most straightforward method is to let the environment itself determine what is fit: if the action maintains the basic organization, it is, otherwise it is not. This can be dangerous, though, since trying out an inadequate action may lead to the destruction of the system. Therefore, complex systems such as organisms or minds have evolved internal models of the environment. This allows them to try out a potential action “virtually”, in the model, and use the model to decide on its fitness. The model functions as a vicarious selector, which internally selects actions acting for, or in anticipation of, external selection. This “shortcut” makes the selection of actions much more reliable and efficient. It must be noted, though, that these models themselves at some stage must have evolved to fit the real environment; otherwise they cannot offer any reliable guidance. Usually, such models are embodied in a separate subsystem, such as the genome or the brain” (Heylighen, 2001).

The sex-stress relationship is nonlinear and is described by approximation as inverted “U”-shaped: sex is favored in intermediate stressful environments, while stable stress-free and extreme stressful environments favor asex (Moore & Jessop, 2003). Constant conditions favor asexuality (Bürger, 1999) which may explain the high incidence of parthenogenesis in environments such as stable forest
soils (Cianciolo & Norton, 2006; Domes et al., 2007). Evolutionary models based on the asexual and sexual replication pathways in Saccharomyces cerevisiae suggested that sexual replication can eliminate genetic variation in a static environment, as well as lead to faster adaptation in a dynamic environment (Gorodetsky & Tannenbaum, 2008).

A change of environmental conditions that reduce Darwinian fitness may increase (i) mutagenesis, (ii) epimutagenesis, (iii) recombination rate, (iv) mutability of simple sequence repeats, and (v) mobilization of transposable elements, all of which, when acting on the germline, increase heritable (epi)genetic variation (Heininger, 2013). Sexual reproduction regulates these processes and, by changing the balance of sexual mutation-selection cascades, modulates the (epi)genetic variation-selection balance. Theoretical models suggest that fluctuating selection is an important factor in maintaining genetic polymorphism (Korol et al., 1996, Kirzhner et al., 1998; Bürger & Gimelfarb, 2002). Likewise, empirical studies of cyclical and fluctuating selection suggest an association between temporal environmental heterogeneity and the amount of genetic variation (Kondrashov & Yampolsky, 1996; Korol et al., 1996).

These environments and their associated stochastic generation of variation appear to have an evolutionary rationale: fighting variation with variation (Ashby, 1956; Meyers & Bull, 2002) creating lottery tickets for the raffle of life. On the other hand, sexual reproduction as evolutionarily highly successful strategy highlights an eminent characteristic of evolution: it pays off to diversify and be prepared for the unlikely event. And: generation of variation is no happenstance outcome but a highly regulated process and environmental stochasticity is its evolutionary “impetus”.

17. Stochasticity and selection: duality in evolution

The paradigm of calculability, determinism and monoausality dominated the sciences until the beginning of the 20th century. Since the end of the 19th century, however, monocausal approaches in many different sciences started to collapse. Even in pure mathematics and logics, problems with the calculability of the universe arose (e.g. Russell’s paradox). Hilbert’s program failed with Kurt Gödel’s proof. At the level of physics, many different problems (e.g. ultraviolet catastrophe, wave-particle duality) led to the development of new physics (Brunner & Klauninger, 2003). Niels Bohr, the “father of quantum mechanics,” indicated that the complementarity predicted and observed in quantum mechanical investigations? such as the wave-particle duality of light and all quanta? was not limited to the quantum realm, but was a more broadly applicable (perhaps universal) concept, which should have correlates in the study of living things (Bohr, 1937; Roll-Hansen, 2000; McKaughan, 2005). Like the wave paradigm could not explain a variety of physical properties of light, explaining evolution by natural selection as only organizing principle has created various implausibilities. As it stands, it is accepted that it makes sense to use stochastic models in population genetics. But why should a selection-only process be stochastic? It is agreed that natural selection has its limits (Barton & Partridge, 2000). But so far these limits have been explained by e.g. genetic architecture, genetic drift, historical contingency or developmental constraints.

Evolution is both the result of random events at all levels of organization of life and of constraints that canalize it, in particular by excluding, by selection, incompatible random explorations. So, ergodic explorations are restricted or prevented both by selection and the history of the organism (Longo et al., 2012). Mayr (2000) wrote: “Darwin settled the several-thousand year-old argument among philosophers over chance or necessity. Change on the earth is the result of both, the first step being dominated by randomness, the second by necessity. only the first step in natural selection, the production of variation, is a matter of chance. The character of the second step, the actual selection, is to be directional.” According to Mayr (1980), selection is “the only direction-giving factor in evolution”. On the other hand, Monod (1971, p. 112-113) argued that “chance alone is at the source of every innovation, of all creation in the biosphere. Pure chance, absolutely free but blind, at the very root of the stupendous edifice of evolution: ...It is today the sole conceivable hypothesis, the only one that squares with observed and tested fact.” Figure 2 depicts the linear evolution model as put forward in the Modern Synthesis (e.g. Mayr, 2000). This linear model contains the elements of selection and chance but lacks a feedback loop and, hence, is unable to learn. Moreover, the model failed to recognize the interaction of stochasticity and selection. The confusion caused by this failure led to the perception of natural selection as a statistical process (see chapter 15).

Darwin already realized that variation is an essential commodity in evolution but he was unaware of its cause. The Modern Synthesis regarded variation as the result of accident, happenstance and imperfection.
The Modern Synthesis draws a non-cybernetic picture of evolution. As outlined by Mayr (see above), stochasticity and natural selection are distinct entities, chance and necessity. Natural selection and random drift can be distinguished from one another (Millstein, 2002; Pfeifer, 2005; but see Abrams, 2007). The cybernetic theory, as advocated here, links both by feedback control: input (environmental stochasticity) determines output (natural selection) and input, at least in part, is determined by output. Importantly, the output signal is a population-level signal including density- and frequency-dependent phenomena that feed back to the individual level of the input signal, inextricably intertwining both stochasticity and natural selection and the individual and population levels of selection. Similarly, in Newtonian mechanics space and time are distinct entities. In Einstein’s Relativity Theory, both are no longer separated but an integrated entity in a four-dimensional continuum of space and time. Intriguingly, the stochasticity-selection duality seems analogous to the wave-particle duality. Schrödinger’s concept of ‘entanglement’ between the states of particles is the key to wave–particle duality (Knight, 1998). ‘Entanglement’, is a peculiar but basic feature of quantum mechanics introduced by Erwin Schrödinger in 1935. Individual quantum-mechanical entities need have no well-defined state; they may instead be involved in collective, correlated (‘entangled’) states with other entities, where only the entire superposition carries information. That may apply to a set of particles, or to two or more properties of a single particle. Likewise, the entangled state of the stochasticity-selection duality can be conceptually disentangled by cybernetic modeling but is phenomenologically an entity.

It is textbook knowledge that selection needs variation to work on. The fundamental question, however, is whether variation is the result of accident and chance or whether it evolved as a means to cover all bases in response to the unpredictability of life. Ashby’s Law of Requisite Variety formulated the conceptual framework to understand how internal variety of a system has to match its external variety. That variation arises at all levels of biological organization such as the genetic, epigenetic, cellular network, developmental, physiological, behavioral and life-history level, that it is malleable in response to stress (when it is most needed) and that sexual reproduction evolved as tool creating pre-selected variation, is evidence Ashby’s Law succinctly describes the cybernetic behavior of evolution.

17.1 The creative conflict between stochastic indeterminism and selective determinism

All processes in Nature are fundamentally stochastic. Poisson tried to model mathematically how one could have stable probabilities of mass phenomena even when the probabilities for individuals are not constant. The law of large numbers teaches that absolute regularity emerges in a long run of draws. He did indeed prove that under certain restrictions, even when the probability at repeated trials is variable, in the long run the average relative frequency does converge on p, the average probability for individual trials (Hacking, 1983) The name “law of large numbers” is still used loosely in probability theory, although there are now so many different theorems that one needs better names, which usually clump around what is called the central limit theorem. The law of large numbers is true for systems at equilibrium, where one can generally expect for a system with N degrees of freedom the relative magnitude of fluctuations to scale as 1/N. However, when the system is driven out of equilibrium, the central limit theorem does not always apply, and even macroscopic systems can exhibit anomalously large (giant) fluctuations (Keizer, 1987; Tsimring, 2014). In the duality of stochasticity and selection, variation is recognized as the result of a multitude of processes, resulting in a bet-hedging response to stochasticity. Ross Ashby’s ‘Law of Requisite Variety’ (1956, p. 206) is the organizing principle of the stochasticity-selection duality. Stochastic environments coerce organisms into lotteries. But today’s winners can be tomorrow’s losers, particularly following natural disaster or epidemic outbreaks. Insurance is a population-level risk-sharing strategy of risk-averse agents buffering against idiosyncratic risk. Via the law of large numbers and bet-hedging, evolution generated a form of automatic biological insurance against idiosyncratic risk (Robson, 1996).

In essence, stochasticity and selection work against each other within the limits of total chaos and complete order, the two extremes where evolution can no longer work. Stochasticity contributes to maladaptation or limits adaptation (Travisano et al., 1995a; Hereford, 2009; Lenormand et al., 2009). On the other hand, stochasticity and selection are interdependent. None can prevail without depriving evolution of its very basis. Selection could not work without the stochastic phenomenon of variation; and stochasticity needs the ordering power of selection to create the complex structures of self-organization (Bak et al., 1987, 1998). Intriguingly, part of the stochasticity is created by selection itself, e.g. through bet-hedging strategies, coevolutionary cycles, density- and frequency-dependent selection, or niche construction (Meyers & Bull, 2002). On the other hand, stochasticity
drives variation and variation is the raw material for selection to work on. Theoretical models suggest that fluctuating selection is an important factor in maintaining genetic polymorphism (Korol et al., 1996, Kirzhner et al., 1998; Bürger & Gimelfarb, 2002). Likewise, empirical studies of cyclical and fluctuating selection suggest an association between temporal environmental heterogeneity and the amount of genetic variation (Kondrashov & Yampolsky, 1996; Korol et al., 1996). Lévy-like search strategies were revealed in analyses of a variety of behaviors from plankton to humans (Viswanathan et al., 1996, 2001; Bartumeus et al., 2003; Barabasi, 2005; Brockmann et al., 2006; Reynolds & Frye, 2007; Reynolds & Rhodes, 2009; Humphries et al., 2010). The models simulating these behaviors combine a multitude of stochastic processes by deterministic rules (Maye et al., 2007).

In addition to the inevitable noise component, a nonlinear signature suggesting deterministic endogenous processes (i.e., an initiator) is involved in generating behavioral variability. It is this combination of chance and necessity that renders individual behavior so notoriously unpredictable (Maye et al., 2007).

Although within wide boundaries, stochasticity and selection have to be balanced. Evolutionary biology already acknowledged mutation-selection equilibrium as evolutionary phenomenon; it is time to realize that there is a stochasticity-selection balance. Too much stochasticity would be detrimental for learning: if the cybernetic feedback concerning fitness effects would not behave with a certain stability and change too irregularly, learning would be impaired. Fortunately, with respect to living organisms, nature is capricious rather than completely random (Lewontin, 1961, 1966). There is a variable degree of ecological predictability: demographic cycles due to e.g. predator/prey interactions, seasons with their cyclicity of resource availability, circadian cycles, tides, etc. Bet-hedging only is favored in an intermediate range of environmental stochasticity. As the environment becomes more stable or more chaotic, bet-hedging strategies have a lower fitness advantage (Philippi & Seger, 1989; Müller et al., 2013). Stochasticity is ambiguous (e.g. beneficial, neutral and deleterious mutations) with regard to outcome while selection filters and directs the ambiguity. And learning attenuates the randomness. Selection is the stabilizing force that brings order into the chaos and provides the feedback for learning to occur. Both stochasticity and selection render evolution opportunistic.

The stochasticity-determinism duality is not adequately reflected by existing models. Modifying Dobzhansky’s notorious quote, Lynch (2007a) wrote: “Nothing in evolution makes sense except in light of population genetics”. However, in 1961 Lewontin did not consider population genetics an “adequate theory of evolutionary dynamics. On the contrary, the theory of population genetics, as complete as it may be in itself, fails to deal with many problems of primary importance for an understanding of evolution.” In this paper, Lewontin (1961) suggested that the modern theory of games (von Neumann & Morgenstern, 1944, 1953) may be useful in finding exact answers to problems of evolution not covered by the theory of population genetics. A first application of game theory to evolutionary issues was the work of Maynard Smith and Price (1973) on animal conflicts and their concept of an “evolutionarily stable strategy” (ESS). The vast body of theoretical work based on the concept of an ESS, however, often disregards environmental stochasticity. For example, Maynard Smith’s often quoted book (Maynard Smith, 1982a) contains no reference to stochasticity. Many models in evolutionary game theory (EGT) involve infinite populations with a deterministic evolutionary dynamic. While these idealizations may provide a good starting point for reasons of mathematical tractability, there are important limitations to them. The methodological focus on equilibria (specifically the ESS) in EGT has resulted in missing important features of evolutionary systems that can only be captured by dynamical analysis (Huttegger & Zollman, 2012). An important feature of EGT models is repetition. If the games were not repeated, these EGT models would not be able to provide any insight into adaptive behaviors and strategies due to the dynamic nature of the mechanisms of evolution. But even standard dynamical analysis has strong idealizations such as infinite populations and deterministic evolution. These kind of idealizations can miss possible explanations, for example regarding the evolution of cooperation (Smead, 2008; Forber & Smead, 2014). Importantly, evolution "plays" both within-generation and trans-generation games. At each game repetition population make-up in turn is determined by the results of all of the previous contests before the present contest- it is a continuous iterative process where the resultant population of the previous contest becomes the input population to the next contest. As stochastic process (Lenormand et al., 2009; Kupiec et al., 2012) evolution can be described by lottery models (Chesson & Warner, 1981; Proulx & Day, 2001; Svardal et al., 2011).

17.1.1 Playing dice with controlled odds

Albert Einstein once said, “I am convinced the Old One [God] does not play dice” (Jammer, 1999, p. 222).
New evidence about the fractal geometry of nature, chaos, and complexity challenges these negative statements about the statistical nature of the physical world (Gleick, 1987). Thus, chaos theorist Joseph Ford remarked: “God plays dice with the universe, but they’re loaded dice” (Gleick, 1987, p. 314).

The indeterminism-determinism interaction is best illustrated by processes at the cellular level. Cell fate decisions are often controlled by both stochastic and deterministic features (Losick & Desplan, 2008; MacArthur et al., 2009; Balázsi et al., 2011; Snijder & Pelkmans, 2011). Thus, genetically homogeneous populations adopt distinct fates? cell fate decisions are stochastic by virtue of the feedback architecture of genetic networks (Smits et al., 2006; Davidson & Surette, 2008) or deterministically linked to the cell cycle or even a combination of both. Examples of cellular-population heterogeneity include differentiation of progenitor hematopoietic stem cells (Mayani et al., 1993), non-genetic individuality in bacterial chemotaxis (Spudich & Koshland, 1976), and epigenetic inheritance and incomplete penetrance of transgenes in mice (Morgan et al., 1999). As a result, bacteria and cells determine their fate by “playing dice with controlled odds” (Ben-Jacob & Schultz, 2010). Constrained randomness, intermediate between rigid determinism and complete disorder is what is usually seen (Theise & Harris, 2006). Specific environmental or genetic cues may bias the process, causing certain cellular fates to be more frequently chosen (as when tossing identically biased coins). Still, the outcome of cellular decision making for individual cells is a priori unknown (Balázsi et al., 2011). Sexual reproduction is the paradigm of this controlled stochastic strategy. The huge (epi)genetic variation that is created by stochastic epimutation and mutagenesis is contained by selection cascades that engender pre-selected variation (Heininger, 2013).

18. The blending of ecology and evolution

In my opinion, the greatest error which I have committed has been not allowing sufficient weight to the direct action of the environment, for example, food and climate, independently of natural selection. When I wrote The Origin, and for some years afterwards, I could find little good evidence of the direct action of the environment; now there is a large body of evidence.
Charles Darwin (1876) in a letter to Moritz Wagner

In most natural populations, the reproductive potential far exceeds the environmental opportunity, and natural selection proceeds by culling to what the habitat can support (King, 1967). As Smith (2012a) put it: “In some respects natural selection is a quite simple theory, arrived at through the logical integration of three propositions (the presence of variation within natural populations, an absolutely limited resources base, and procreation capacities exceeding mere replacement numbers) whose individual truths can hardly be denied.” The resulting struggle for existence is the engine that drives evolution. Haeckel (1866) defined ecology as the science of the struggle for existence (Cooper, 2003). Thus, from early on, ecology and evolution have been intertwined. In this vein of thought Van Valen (1973b) described evolution as “the control of development by ecology”. Calls for an ‘integrative’ understanding of biological processes keep being repeated in the literature, from Dobzhansky’s (1973) famous quote “Nothing in biology makes sense except in the light of evolution” to current, more focused statements that evolution itself only makes sense when viewed in its ecological context (Coulson et al., 2006; Saccheri & Hanski, 2006; Johnson & Stinchcombe, 2007; Metcalf & Pavard, 2007; Pelletier et al., 2007; 2009; Kokko & López-Sepulcre, 2007; Blute, 2008; Grant & Grant, 2008; Bassar et al., 2010; Matthews et al., 2011; Schoener, 2011). The repeated call for an integrative view of ecology and evolution only reflects the still existing division between ecology and evolution despite Grant and Grant's (2008) dictum: “Nothing in evolutionary biology makes sense except in the light of ecology.” Notwithstanding some recent relevant studies, the importance of the evolution-to-ecology pathway across systems is still considered unknown (Schoener, 2011). The feedbacks between ecological and evolutionary changes are now known to be bidirectional (Post & Palkovacs, 2009; Schoener, 2011; Miner et al., 2012). Thus, “Nothing in biology makes sense except in the light of an integrated perspective of both ecology and evolution”. The stochasticity-natural selection duality finally blends ecology and evolution into each other. Einstein introduced the concept of space-time as a single entity. Stochasticity and natural selection interact on a variety of levels (Abrams, 2007), and, in fact, form a single entity.

19. Cutting the Gordian knot of controversies

Evolutionary theory is the arena of a multitude of controversies. Particularly, the level of selection issue and sociobiology have seen rancorous theoretical
debates.
In chapter 15 it has been argued that environmental stochasticity changes the rules of evolution. Darwinian tradition with its assumption of constant environments emphasizes the role of individual selection. Stochastic environments coerce individuals into lotteries. The risk-averse individuals, on the other hand, employ the risk-sharing strategy of insurance. Risk-sharing can only be done in groups, the larger the better. Thus, stochastic environments turn individual selection into multilevel selection.

At the heart of the debate in sociobiology is how cooperation and altruism can persist in the face of cheating (Wade & Breden, 1980; Hamilton & Taborsky, 2005; Bijma et al., 2007). Some have suggested that the solution to this problem is the level of selection (Slatkin & Wade, 1978; Wade, 1978a; Wilson & Sober, 1994; Keller, 1999; Goodnight, 2005; Wilson, 2005; Nowak et al., 2010). In both biology and the human sciences, social groups are sometimes treated as adaptive units, whose organization cannot be reduced to the individual level. In this view, group-level adaptations can evolve only by a process of natural selection acting at the group level (Wilson & Sober, 1994). This group-level view is opposed by a more individualistic one that treats social organization as a by-product of self-interest, suggesting that altruism can evolve through individual selection depending on the degree of relatedness within a group (Hamilton, 1964; Wade, 1978b; Michod, 1982). More recent approaches treat multilevel selection as a continuum, in which fitnesses of individuals depend on both individual and group properties, of which pure group selection and individual selection are limiting cases (Keller, 1999). I elaborated a theory explaining the ecology-driven pattern of social interactions based on the insight that environmental stochasticity favors the evolution of cooperation as bet-hedging behavior (Heininger, 2015).

Another conundrum of evolutionary biology and population genetics is the coexistence of two basic observations (Walsh & Blows, 2009; Jeffreys et al., 2012): in natural populations genetic variation is found in almost all traits (Mousseau & Roff, 1987; Houle, 1991, 1992, 1998; Hill & Caballero, 1992; Lynch & Walsh, 1998) in the presence of strong stabilizing selection (Haldane, 1949; Clarke, 1979; Endler, 1986; Kingsolver et al., 2001; Hereford et al., 2004; Johnson & Barton, 2005). These two observations are in direct conflict as stabilizing selection should deplete genetic variation (Bürger & Gimelfarb, 1999; Tomkins et al., 2004; Johnson & Barton, 2005; Walsh & Blows, 2009). In general, maintenance of genetic variation is linked with environmental heterogeneity (Hedrick, 1986; Futuyma & Moreno, 1988; Wilson, 1994; MacDonald, 1995; Ellis et al., 2006). Thus, genetic variation in populations is the evolutionary footprint of temporally and spatially stochastic environments (Antonovics, 1971; Gillespie, 1973b; Hedrick et al., 1976; Hedrick, 1986; Mitchell-Olds, 1995; Sasaki & Ellner, 1995, 1997; Ellner, 1996; Bürger & Gimelfarb, 2002; Leimar, 2005; Heininger, 2013). Moreover, the biodiversity of species is mainly supported by habitat heterogeneity and niche partitioning. As a result, a positive relationship between species richness and habitat heterogeneity is predicted (Hutchinson, 1957; MacArthur, 1972; Petren, 2001; Kallimanis et al., 2008; de Souza Júnior et al., 2014).

A variety of other evolutionary controversies and conundrums can be resolved by the stochastic environment paradigm as was discussed in Heininger (2013).

20. Abbreviations

EGT: evolutionary game theory
ESS: evolutionarily stable strategy
SOC: self-organized criticality

21. References


Anholt BR, Werner EE (1995) Interaction between food availability and predation mortality mediated by


Biro PA, Post JR, Parkinson EA (2003b) Population


Boyce MS (1979) Seasonality and patterns of natural


Caraco T (1983) White-crowned sparrows (Zonotrichia...


Dubnau D, Lovett CM Jr (2002) Transformation and...


Ellegren H (2004) Microsatellites: simple sequences...


Falster DS (1990) Selection in different environments: effects on environmental sensitivity (reaction norm) and on mean performance. Genet Res 56: 57–70.


Frank SA (1996b) Models of parasite virulence. Q Rev
Biol 71: 37–78.


Fujita M, Losick R (2005) Evidence that entry into sporulation in Bacillus subtilis is governed by a gradual increase in the level and activity of the master regulator Spo0A. Genes Dev 19: 2236–2244.


Goodnight CJ (1990a) Experimental studies of...


Springer.


Hooshangi S, Bentley WE (2008) From unicellular properties to multicellular behavior: bacteria quorum


383–384.


http://evolutionlist.blogspot.com/2009/03/are-mechanisms-that-produce-phenotypic.html


McKaughan DJ (2005) The influence of Niels Bohr on


Moraiti CA, Nakas CT, Papadopoulos NT (2012)


Parker GA, Immler S, Pitnick S, Birkhead TR (2010) Sperm competition games: sperm size (mass) and
Pfeifer J (2005) Why selection and drift might be


Richardson K (2012) Heritability lost; intelligence found. Intelligence is integral to the adaptation and survival of all organisms faced with changing environments. EMBO Rep 13: 591–595.


Rosenfeld S (2011) Mathematical descriptions of...


Siegl-Cachedenier I, Munoz P, Flores JM, Klatt P,


Smith SE, Au DW, Show C (1998) Intrinsic rebound


Spalding DA (1837) Instinct with original observations on young animals. Macmillian’s Magazine 27: 282–293.


Tebbich S, Teschke I (2014) Coping with uncertainty: Woodpecker finches (Cactospiza pallida) from an unpredictable habitat are more flexible than birds from a stable habitat. PLoS ONE 9: e91718.


Wade MJ (1978b) Kin selection: a classical approach and a general solution. Proc Natl Acad Sci USA 75:


Weber M (2001) Determinism, realism, and probability


Weinreich DM, Chao L (2005) Rapid evolutionary escape by large populations from local fitness peaks is likely in nature. Evolution 59: 1175–1182.


Wright S (1932) The roles of mutation, inbreeding,


Zeineddine M, Jansen VAA (2009) To age, to die:


Illustrations

Illustration 1

Fig 1

A. Cybernetic systems are characterized by feedback control. They are a special class of cause-and-effect (input-output) systems. Learning automata are adaptive decision-making devices operating on unknown random environments. Iteration is required for feedback control.

B. In artificial selection, the breeder determines the constant direction of the breeding goal and selects the individuals for the next round of breeding.

C. In evolution, the direction and regime of selection are established by the environment. However, the target, adaptation to varying biotic and abiotic environmental conditions, is a moving target and selection can be highly fluctuating. Bet-hedging is the adaptive response to environmental stochasticity. The feedback between output and input signals inextricably intertwines both stochasticity and natural selection.

Illustration 2

Fig 2
Figure 2

The linear evolutionary theory as put forward in the Modern Synthesis (e.g. Mayr, 2000). Although this linear model contains the components of selection and chance it lacks a feedback loop (see figure 1C) and, hence, is unable to learn.