# The Fundamental Problem with the Building Block Hypothesis

**Keki M. Burjorjee**                                         kekib@cs.brandeis.edu
Department of Computer Science, Brandies University, Waltham, MA 02454, USA

**Abstract**

Skepticism of the building block hypothesis has previously been expressed on account of the weak theoretical foundations of this hypothesis and anomalies in the empirical record of the simple genetic algorithm. In this paper we focus on a more fundamental cause for skepticism—the extraordinary strength of some of the assumptions undergirding the building block hypothesis. As many of these assumptions have been embraced by the designers of so called "competent" genetic algorithms, our critique is relevant to an appraisal of such algorithms. We argue that these assumptions are too strong to be acceptable without additional evidence. We then point out weaknesses in the arguments that have been provided in lieu of such evidence.

**Keywords**

Genetic algorithms, building block hypothesis, philosophical foundations, occam's razor, population genetics.

## 1   Introduction

In constructing a representation (a genome-to-phenotype map and a fitness function) a GA practitioner implicitly determines how fitness gets distributed over a genome set. If a GA with the representation is adaptive then, with overwhelmingly high probability, the induced fitness distribution has some type of "structure" that the GA is exploiting. There can be no other reason for the GA's performance. GAs are frequently adaptive in practice. This entails that GA practitioners often construct representations that induce fitness distributions with GA-exploitable structure.

Before proceeding further, let us clarify what we mean by the word "adaptive". Given a search problem, let us say that some population based search algorithm is *adaptive* if, across several runs with different random seeds, the average fitness of the population consistently trends upwards. By this token GAs are often adaptive in practice, whereas population based random search is not. We refer to this feature of genetic algorithms as their *adaptive capacity*.

We posit that *any* coherent theory about the adaptive capacity of genetic algorithms must consist of the following: Firstly a set of assumptions about the way fitness commonly gets distributed via the representational choices of GA practitioners; let us call these the *fitness distribution assumptions*. And secondly, a hypothesis about how the assumed "structure" in the fitness distributions gets exploited by a GA during adaptation; let us call this the *exploitation hypothesis*.

The exploitation hypothesis depends critically on the fitness distribution assumptions, but not vice-versa. The fitness distribution assumptions, in other words, are

foundational. Therefore, when developing an explanation for the adaptive capacity of genetic algorithms, getting the fitness distribution assumptions right is extremely important. Fundamentally flawed fitness distribution assumptions thwart the entire enterprise no matter how much effort is lavished upon the development and justification of the exploitation hypothesis.

To be scientifically viable, an exploitation hypothesis must be based upon fitness distribution assumptions that are *weak*. This is nothing but an application of the principle of Occam's razor, which holds that the weaker the assumptions that undergird a theory, the more viable the theory. This principle clearly makes sense in the current context. GA representations are rather ad-hoc. Therefore the weaker our assumptions about the structure of the the induced fitness distributions, the more likely it is that these assumptions hold true[1].

The building block hypothesis (BBH) is currently the dominant explanation for the adaptive capacity of simple genetic algorithms. Though this hypothesis has come in for some sharp criticism in recent times, it remains the compass by which the vast majority of GA practitioners—past and present—have allowed themselves to be guided. It is also the first explanation for the adaptive capacity of GAs that most students receive, and in this capacity probably exerts an *anchoring effect* (Tversky and Kahneman, 1974) on their reasoning.

Previously expressed skepticism of the building block hypothesis can be divided into two categories. The first consists of criticism of the weak theoretical foundations of this hypothesis (for a survey see Reeves and Rowe, 2003, section 3.3). Proponents of the building block hypothesis have, for the most, part brushed aside criticism of this sort. Goldberg, for example, calls such critiques "[a] favorite parlor game in genetic and evolutionary computation circles" (Goldberg, 2002, p7), and, by way of analogy, characterizes such concerns as absurd. "[T]he very idea that an airplane is ineffective or unsafe because a formal mathematical proof of flight does not exist is itself an absurdity.", he writes. "Yet, if this is so—and no proof does exist of airplane flight—and if I ...transform the aircraft into a *genetic algorithm. . .*, why is it that [this] patently absurd alarm seems so real in the context of GAs and their design and use?" (Goldberg, 2002, p19).

The second category is comprised of skepticism arising from the anomalous performance of the simple genetic algorithm on some basic empirical tests (Forrest and Mitchell, 1993), (Watson, 2006, Section 6.2). In response to these results proponents of the building block hypothesis have downplayed the importance of the simple genetic algorithm (Holland, 2000; Goldberg, 2002), and have advocated the use other sorts of genetic algorithms (e.g. cohort genetic algorithms, "competent" genetic algorithms), that are more complicated than the simple genetic algorithm, and typically contain mechanisms that are not biologically plausible. Strictly speaking the building block hypothesis applies only to the simple genetic algorithm. Therefore by downgrading the importance of the simple genetic algorithm proponents of the building block hy-

---

[1] Adherents of the building block hypothesis may disagree with our assessment that the representations they construct are "rather ad-hoc". After all, much advice for constructing representations has been dispensed (e.g. "ensure a large supply of building blocks"); those who have made an effort to follow this advice may claim to have a basis for making strong assumptions about the structure of the fitness distributions they induce. Unfortunately claims of this nature are unjustified. While there is plenty of advice on how representations should be constructed, there is, as far as we can tell, no principled way for determining how this advice should be put into practice in specific instances (except perhaps when the problems are contrived). In this respect the dispensed advice is much like the famous investing mantra "buy low, sell high"—easy to state, hard to implement.

pothesis can claim to have rendered skepticism about the veracity of this hypothesis irrelevant.

The skepticism expressed in this paper does not belong to either of the two categories mentioned above. It is, in a sense, more *fundamental*, stemming from a critical appraisal of the strength of the fitness distribution assumptions that undergird the building block hypothesis. We examine various influences—both historical, and social—that have shaped these assumptions, and argue that the assumptions are too strong to be acceptable in the absence of empirical evidence. We then examine the kinds of arguments that proponents of the building block hypothesis have provided in lieu of such evidence. As the assumptions undergirding the building block hypothesis have largely been embraced by the designers of the new types of algorithms mentioned in the previous paragraph, our criticism is highly relevant to an appraisal of these algorithms.

The rest of this paper is organized as follows. In section 2 we briefly recount the history of the building block hypothesis—its origin, ascent, and recent troubles. In section 3 we describe the fitness distribution assumptions undergirding this hypothesis, and explain why we find them to be unacceptably strong. In sections 4 and 5 we critically examine the ways in which proponents of the building block hypothesis have sought to justify this hypothesis, and by extension, the fitness distribution assumptions that undergird it. In section 6 we explain the negative ramifications of our criticism for a popular line of research within the genetic algorithmics community. And finally, in our conclusion, we draw parallels between the fitness distribution assumptions of the building block hypothesis, and the now defunct concept of luminiferous aether that was popular in nineteenth century theories about the propagation of light.

## 2 A Brief History of the BBH

Scientific theories are typically presented without reference to the context within which they were developed (Okasha, 2002, p79). At a certain level this of course makes sense; surely what a scientist has for breakfast is immaterial to one's evaluation of her theories. At the same time, it has been observed that to genuinely understand the state of a science, an acquaintance with its history is essential. Every scientific theory is undergirded by what the historian of science, Thomas Kuhn, calls *received beliefs* (Kuhn, p4)—assumptions transmitted from one generation to the next within a scientific community. An acquaintance with the history of a science can help one identify the origins of such assumptions and the circumstances under which they have been perpetuated.

### 2.1 Origin of the Building Block Hypothesis

In the 1960s and early '70s Holland developed an abstract mathematical model of adaptive processes (what he called an *adaptive plan*) which was inspired by natural evolution but was of greater generality. Holland sought to use this model to unify under one theoretical framework adaptation in such diverse fields as neuroscience, economics, control, game theory, artificial intelligence, and genetics.

For any adaptive process that generates structures of some type, one can define a set $\mathcal{A}$ of all the structures that may be generated during the adaptive process. For example, in the domain of economic planning an element of $\mathcal{A}$ may be a mix of goods; in game-theory $\mathcal{A}$ may be the set of all strategies with respect to some game (Holland, 1975, p4). Holland conceptualized adaptation as a process that starts out by sampling from all of $\mathcal{A}$, and, over time, starts restricting its sampling, to subsets of $\mathcal{A}$ with in-

creasing average fitness.

The central conceptual objects in Holland's framework are particular subsets of $\mathcal{A}$ called schemata (singular schema). Holland noted that each point in $\mathcal{A}$ may belong simultaneously to several schemata. Therefore an evaluation of the fitness of that point is, in effect, a fitness evaluation of a sample from each schema that the point belongs to. If the point is 'fit' then this reflects well on all of those schemata, and vice-versa if the point is 'unfit'. This observation suggested to Holland the possibility of the existence of algorithms which, by testing small numbers of points, implicitly test vast numbers of schemata, and then implicitly use this information to concentrate trials in schemata of increasing average fitness. Holland named this phenomenon *intrinsic parallelism* (Holland, 1975, p74), a name he later revised to *implicit parallelism* (Holland, 1992). Holland also clearly seems to have been impressed by the utility of *hierarchical* assembly (Simon, 1969, Chapter 4), (Holland, 1975, p168).

In a long argument that freely mixed speculation and deductive reasoning—the line between the two was typically left blurry—Holland concluded that a specific adaptive plan that models natural evolution—what he called a *genetic plan*—will generate high-fitness solutions to difficult adaptation problems, and that this adaptive plan will do so using implicit parallelism and hierarchical assembly (Holland, 1975).

Starting in 1970, Holland's students began applying implementations of genetic plans to adaptation problems (e.g. Cavicchio, 1970; Hollstien, 1971). They found that these algorithms typically outperformed random search. In a landmark dissertation, De Jong (1975)described experiments in which a stripped down version of the genetic plan—what is now called the simple genetic algorithm—was applied to a carefully contrived set of fitness functions with well-understood and diverse characteristics. De Jong reports that "Out of these early studies [his and those of his colleagues] emerged a picture of a GA as a robust adaptive search procedure, which was surprisingly effective as a global search heuristic" (De Jong, 2006).

Holland's theoretical work on genetic plans was seen as the obvious place to look for an explanation for the adaptive capacity of genetic algorithms. Holland and his students simplified this work and settled upon the well known explanation that goes by the name of the building block hypothesis (Goldberg, 1989; Holland, 1992; Mitchell, 1996).

## 2.2   Initial Espousal and Recent Skepticism

Until the late 1980s the building block hypothesis seems to have gone relatively unquestioned. However, with the explosive increase in the popularity of the genetic algorithm came increasing scrutiny of its theoretical foundations. Starting in the early 1990s several researchers began to publish independent ground-up theoretical analyses of the dynamics of genetic algorithms (Vose and Liepins, 1991; Nix and Vose, 1992; Vose, 1993; Prügel-Bennett and Shapiro, 1994; Rattray, 1996; Shapiro, 2001). What prompted these entirely new lines of theoretical analysis? It is hard to say for certain, but we believe that at least part of the cause was frustration with the unclear demarcation between speculation and mathematically provable facts that is characteristic of the argument for the building block hypothesis. In the preface to his book on the simple genetic algorithm, Vose writes "My central purpose in writing this book is to provide an introduction to what is known about the theory of the Simple Genetic Algorithm. The rigor of mathematics is employed so as not to inadvertently repeat myths or recount folklore" (Vose, 1999). He adds that the absence of core elements of "standard GA theory" in his book is

due to the unintelligibility, the irrelevance, or the mathematical unjustifyability of these elements. In a later work Wright et. al. remark, "The various claims about GAs that are traditionally made under the name of the *building block hypothesis* have, to date, no basis in theory, and in some cases, are simply incoherent" (Wright et al., 2003). Fogel (2000, p117, 169) has expressed similar sentiments.

Statements like these have served a vital purpose. Despite it's name, the building block hypothesis had come to be treated as much more than a hypothesis. It had become the *de-facto* explanation for the success of genetic algorithms, thoroughly shaping the paradigm within which most GA research was conducted. For example, this hypothesis determined what constituted a valid question, a valid explanation, a valid prediction, and a valid enhancement of some genetic algorithm. Given the assertive tone in which the building block used to be presented (see Goldberg, 1989; Holland, 1992), and the blurry line between deductive reasoning and speculation in the argument for this hypothesis, students and non-theoreticians eager for an explanation cannot not be blamed for surmising that the building block hypothesis is based largely on deductive reasoning. Statements such as the ones reproduced above now serve to caution them against this notion. Such statements have also served as a call to apologists for this hypothesis to clearly describe their premises and modes of reasoning. The responses elicited (e.g. Holland, 2000; Goldberg, 2002), provide us with the clearest picture yet of the presumptions undergirding the building block hypothesis.

## 3   The Fitness Distribution Assumptions Undergirding the BBH

Let us quickly recount some basic elements of schema theory. In the case of genetic algorithms, $\mathcal{A}$ is the set of all strings of some predetermined length drawn from some alphabet (in what follows we assume that this alphabet is {0,1}). Let us call the elements of this set *genomes*. Schemata are represented by so called 'similarity templates'. Suppose $\mathcal{A}$ is the set of all strings of length 6, then the schema 1*0**0 is the subset of strings in $\mathcal{A}$ with 1 in the first position, zero in the third and sixth positions, and either 1 or 0 in the second, fourth, and fifth positions; the symbol *, called a 'wildcard', stands for 'don't care'. For the sake of brevity, a schema template is often just called a schema. It is important, though, to keep in mind the distinction between the two. Given some population, the *frequency* of a schema is the number of genomes in the population that belong to that schema. The *defining length* of a schema is the difference between the indices of the last and first non-wildcard loci. Finally, the *order* of a schema is the number of non-wildcard loci. Thus, the defining length and order of the schema in the example above are five and three respectively. A schema with low defining-length (and therefore low order) is said to be 'short'.

Let $S_1$ and $S_2$ be two subsets of $\mathcal{A}$. For a population of size $N$, we say that the sampling fitness of $S_1$ is greater than (or less than) the sampling fitness of $S_2$ if the the average fitness of $N$ samples drawn from the uniform distribution over $S_1$ is greater than (or, respectively, less than) the average fitness of $N$ samples drawn from the uniform distribution over $S_2$.

Given some collection of subsets of $\mathcal{A}$ with a non-empty intersection, we say that the intersection is *antagonistic* if there is a high probability that the sampling fitness of the intersection will be less than the sampling fitness values of all the intersecting subsets (Figure 1a). We say that the intersection is *synergistic* if there is a high probability that the sampling fitness of the intersection will be greater than the sampling fitness values of all the intersecting subsets (Figure 1b) . If the intersection is neither

5

antagonistic, nor synergistic, we say that it is *consonant* (Figure 1c)

We define a *basic building block* to be a short schema with sampling fitness that is likely to be greater than the sampling fitness of $\mathcal{A}$. A synergistic intersection between a small collection of basic building blocks is called a 2nd level building block, a synergistic intersection between a small collection of 2nd level building blocks is called a 3rd level building block, and so on. The building block hypothesis rests on assumptions that are *at least as strong as* the following two assumptions:

- **Abundant Basic Building Blocks:** A large number of basic building blocks exist.

- **Heirarchical Synergism:** Antagonistic intersections between collections of building blocks at the same level are rare, whereas synergistic intersections between small collections of co-located building blocks at the same level are common.

We trust that the reader, upon seeing these assumptions explicitly laid out, will agree that they are too strong to be acceptable without additional evidence. Remember that we are talking about the structure of fitness functions induced by the ad-hoc decisions of GA practitioners engaged in solving poorly understood problems.

The most persuasive evidence, of course, would be a body of experiments in which these assumptions are tested for, and found to hold in fitness distributions arising in practice (i.e. ones in which the assumptions have not been implemented *a priori*). Such evidence has not been forthcoming. Instead, proponents of the building block hypothesis have attempted to justify their belief in these assumptions through other means—by appealing to certain metaphysical positions, and by appealing to authority. We now consider each appeal in turn.

## 4 Appeal to Metaphysical Positions

In support of the building block hypothesis Holland has asserted the *building block thesis*. He describes this thesis as follows: 'The "building block thesis" holds that most of what we know about the world pivots on descriptions and mechanisms constructed from elementary building blocks'. He characterizes building blocks as "parts" such that "(i) they must be easy to identify (once they've been discovered or picked out), and (ii) they must be readily recombined to form a wide variety of structures (much as can be done with children's building blocks)".

The way the building block thesis is phrased it is a statement about what is already known about the world rather than a universal law, which is how Holland goes on to use it. Our most sympathetic rephrasing of the building block thesis in light of the way Holland uses it to support the building block hypothesisis is as follows:

Building blocks play a key role in the structure and function of *most* objects and processes. Building blocks are (i) parts of wholes, (ii) easily identifiable, and (iii) recombinable with other building blocks to form a wide variety of forms.

It is necessary to quote Holland at length so as not to leave out any part of his argument for this thesis. Holland writes:
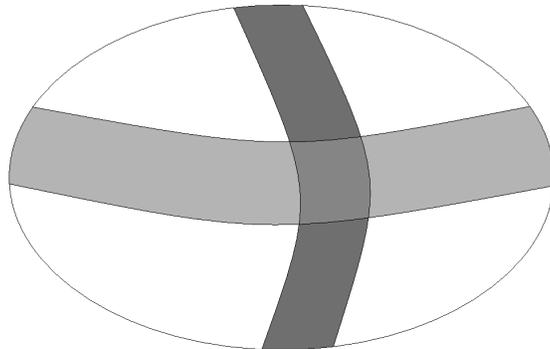
'The successive levels of building blocks used in physics are familiar to anyone interested in science—nucleons constructed from quarks, nuclei constructed

6
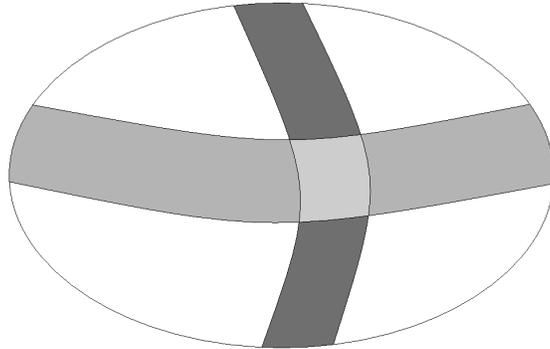
(a) An antagonistic intersection between two subsets
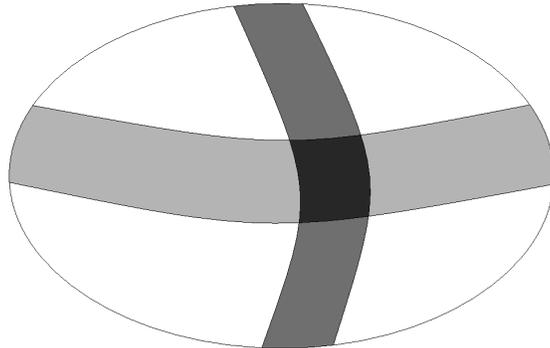


(b) A synergistic intersection between two subsets



(c) A consonant intersection between two subsets

Figure 1: A depiction of antagonistic, synergistic, and consonant intersections between two subsets. Darker shades depict greater sampling fitness

from nucleons, atoms constructed from nuclei, molecules constructed from atoms, and so on ... Nowadays a similar succession presents itself in daily newspaper articles discussing progress in biology: chromosomal DNA constructed from 4 nucleotide building blocks; the basic structural components of enzymes: alpha helices, beta sheets, and the like, constructed from 20 amino acids; standard "signalling" proteins for turning genes "on" and "off," and "autocatalytic bio-circuits," such as the citric acid cycle, that perform similar functions over extraordinarily wide ranges of species organelles constructed from situated bio-circuits, and so on ... And, of course, there are the long-standing taxonomic categories: species, genus, family, etc., specified in terms of morphological and chromosomal building blocks held in common. However the pervasiveness of building blocks only becomes apparent when we start looking at other areas of human endeavor. In some cases we take building blocks so much for granted that we're not even aware of them. Human perception is a case in point. The objects we recognize in the world are always defined in terms of elementary, reusable building blocks, be they trees (leaves, branches, trunks, ...), horses (legs, body, neck, head, blunt teeth, ...), speech (a limited set of basic sounds called phonemes), or written language (the 26 letters of English, for example).

'In other cases we just don't make the building blocks explicit. Consider two major inventions of the 20th century, the internal combustion engine and the electronic computer. The building blocks of the internal combustion engine: gears, Venturi's aspirator, Galvani's sparking device, and so on, were well-known prior to the invention. The invention consisted in combining them in a new way. Similarly, the components of early electronic programmable computers: wires, Geiger's counting device, cathode ray tubes, and the like, were well-known. Even earlier, Babbage had spelled out an overall architecture using long-standard, mechanical building blocks, (gears, ratchets, levers, etc.). The latter invention consisted in combining the electronic building blocks in a way that implemented Babbage's mechanical layout. And, of course, building blocks underpin the critical step for universal computation: arbitrary algorithms are constructed by combining copies of a small set of basic instructions. For both the internal combustion engine and the programmable computer, the building blocks were a necessary precursor, but the innovation required a new combination of the blocks' (Holland, 2000).

Holland regards the *schema theorem* as the Rosetta stone that shows how the building block thesis applies in the specific case of the genetic algorithm (Holland, 2000). This theorem shows that if a short schema with frequency $x$ has above average fitness in some generation $t$, then the expected frequency of that schema in generation $t + 1$ is greater than $x$. From this result Holland concludes that short schemata with above average fitness are the basic building blocks that genetic algorithms implicitly use. That such building blocks must "therefore" be abundant and hierarchically synergistic presumably "follows" from the building block thesis.

The problem with the building block "thesis" is that it is too vague to be *falsifiable*. Firstly, it uses highly subjective language. The stipulations that building blocks be "easy" to identify and "recombinable" beg the question "according to whom?". Secondly, the use of the word "most" makes this "thesis" impossible to falsify unless one conducts an inventory of all entities in the universe.

8

The falsifiability criterion was formulated by the philosopher of science Karl Popper (Popper, 2007b,a) as a way to distinguish between *scientific* theories, such as Einstein's theory of gravitation, and *pseudo-scientific* theories, such as astrology, or Freud's theory of psycho-analysis. Pseudo-scientific theories, noted Popper,

> "appeared to be able to explain practically everything that happened within the fields to which they referred. The study of any of them seemed to have the effect of an intellectual conversion or revelation, opening your eyes to a new truth hidden from those not yet initiated. Once your eyes were thus opened you saw confirming instances everywhere: the world was full of *verifications* of the theory. Whatever happened always confirmed it. Thus its truth appeared manifest and unbelievers were clearly people who did not want to see the manifest truth" (Popper, 2007a, p45).

Scientific theories, in contrast, are theories that take risks—by predicting unexpected phenomena (e.g. gravitational lensing) they leave themselves open to refutation. "A theory which is not refutable by any conceivable event is non-scientific", wrote Popper. "Irrefutability is not a virtue (as people often think) but a vice" (Popper, 2007a, p46).

Consider for example the difference between the Church-Turing thesis (Copeland, 2004), a refutable and therefore scientific thesis, and the building block "thesis", which is neither. Perhaps the most generous way to regard Holland's "thesis" is as a *metaphysical* theory of pan-modularity and pan-hierarchism. Should this new view of the building block thesis allay our concerns about the building block hypothesis? It should not. Indeed *any* hypothesis can be justified by asserting a generalization of the hypothesis as a new metaphysical position, and then using the metaphysical position to argue in favor of the specific hypothesis.

## 5   Appeal to Authority

It is clear from his writings that Holland views the building block hypothesis as a straightforward generalization of Fisher's theory of adaption (Holland, 1975, p89); (Holland, 2000)—a generalization indeed that, to Holland's mind, rests on assumptions that are much *weaker* than Fisher's. Fisher's theory currently reigns as the orthodox view in population genetics—the theoretical branch of evolutionary biology—so it is understandable that Holland finds it exasperating (Holland, 2000) that the building block hypothesis should meet with the kind of criticism it has received from certain quarters within the genetic algorithmics community.

In this section we describe some of Fisher's assumptions, highlight their extraordinary strength, and examine the circumstances under which these assumptions became part of the orthodoxy of population genetics. It is important to stress that the adoption of Fisher's assumptions by population geneticists, though pervasive, is by no means unanimous. Accordingly, we will review some of the criticism that has been leveled at these assumptions. Finally we compare Holland's assumptions with Fisher's. We argue that the extent by which the former are weaker than the latter has been exaggerated, and explain how Holland's assumptions are in fact stronger than Fisher's in an important respect.

## 5.1 The Fisherian Pardigm and its Discontents

For almost two decades after the discovery, in 1900, of his paper on inheritance in pea plants, Mendel's theory of particulate inheritance was thought to be at odds with the theory of adaptation by natural selection proposed by Darwin and Wallace. The *Mendelians* argued that new species arise not by gradual changes, but by large jumps— called saltations—caused by macromutations. Such mutations (which were assumed to occur infrequently) were thought to be the main driver of adaptation. Natural selection on the other hand was thought to play at best a minor part—that of mopping up deleterious macromutations. Though this point of view was never articulated by Mendel himself, his name became associated with researchers such as DeVries, Bateson and Johansen who used the results of Mendel's paper to downplay the effects of natural selection (Mayr, 2003).

Opposing them, the *biometricians* (e.g. Pearson, and Weldon) noted that gradualness abounds in nature, and argued that evolution consists of a gradual shift of an entire population rather than the creation of new types by macromutation. The biometricians made extensive use of statistics to study the effect of natural selection on phenotypic distributions and claimed superiority over the Mendelians on account of their commitment to mathematical rigor (the Mendelians in turn trumpeted their fidelity to empiricism). By and large the biometricians rejected Mendelian inheritance (Provine, 2001, p85). This may seem odd in this day and age, but remember that we are discussing a time that predates the discovery of the material basis of inheritance (chromosomes comprised of DNA), as well as the mechanisms of inheritance (e.g. meiosis).

The possibility of reconciling the views of the Mendelians with those of the biometricians had been considered by Yule, and Parson as early as 1902. However it is Fisher's paper, published in 1918, on "The correlation of relatives on the supposition of Mendelian inheritance" that is widely regarded as marking the beginning of the synthesis of these two purportedly irreconcilable theories into a single theory of evolution—what we now call the *modern synthesis*.

In this paper, Fisher, himself an eminent statistician, presented a mathematical model that incorporated natural selection and Mendelian inheritance, and used it, amongst other things, to calculate correlations between certain traits in relatives. He argued that his model could account for published biometric data (Pearson and Lee, 1903) which Pearson (a biometrician) had previously used to question the adequacy of Mendelian inheritance.

Provine reports that "Fisher's 1918 paper was well received by the few geneticists who could understand his mathematics" (Provine, 2001, p147). We can reasonably suspect that Fisher's use of Pearson's own data to challenge Pearson's position must have been viewed as a coup d'état of sorts. At any rate, there seems to have been no alarm over the remarkably strong assumptions undergirding Fisher's work. What were these assumptions?

In Fisher's model a quantitative trait (e.g. stature, i.e. height) is under the influence of multiple genes, each with multiple allelic instantiations. Fisher assumed that the effects of allele substitutions in an individual were constant and combined additively. This assumption entails that the substitution of one allele for another always has the same additive effect on the value of the trait, regardless of the genetic background in which the allele substitution occurs. To be fair Fisher did discuss the possibility that the effects of allele substitutions might not combine additively. He called this condition

*epistacy*, a term he later revised to *epistasis*. Early in the paper Fisher urged his readers to treat epistasis and the effects of the environment as one might regard "an arbitrary error introduced into the measurements", in other words, as noise. Later he returned to show how the case when epistasis is not well modeled by noise might be dealt with. Unfortunately, his treatment was rather incomplete; he limited himself to addressing "deviations from linearity as may exist between two factors", but not more, because (amongst other things) he deemed such higher order deviations "improbable".

Importantly, considerations of epistasis did not figure into Fisher's accounting for the biometric data of Pearson and Lee. Fisher, in other words, accounted for this data while making very strong assumptions about the effects of allele substitutions. Nevertheless, he seemed sure that his approach was valid. In the conclusion of the paper he wrote:

> "Throughout this work it has been necessary not to include any avoidable complications, and for this reason the possibilities of Epistacy have only been touched upon, and small quantities of the second order have been steadily ignored. In spite of this, it is believed that the statistical properties of any features determined by a large number of Mendelian factors have been successfully elucidated".

Fisher's characterization of epistasis as an "avoidable complication" to his theory betrays a confusion about the role of parsimony in scientific theorizing. He seems to have believed that a simpler *explanation* is to be preferred to a more complicated one. Occam's razor, or the principle of parsimony, however, is applicable to the assumptions undergirding an explanation, *not* to the explanation itself. To see this clearly, note that *any* phenomenon can be succinctly accounted for by the short explanation, "It is the will of God". The appeal of scientific theories lies, certainly not in their comparative succinctness (scientific explanations are always longer), but in the comparative parsimony of the *assumptions* involved.

As mentioned above, the extraordinary strength of Fisher's assumptions drew no protest at the time. Encouraged by the reception of his paper, Fisher continued with his particular mode of analysis (Provine, 2001, p147). In his highly influential book, in which he set Darwin's theory of natural selection on a Mendelian foundation, Fisher (1930) gave short shrift to epistasis.

This time however, Fisher's assumptions were criticized by Sewall Wright, who, based on years of experimental work, was convinced of the error of assigning fitness effects to individual genes. In a review of Fisher's book, Wright remarked that the Fisherian approach

> "assumes that each gene is assigned a constant value, measuring its contribution to the character of the individual (here fitness) in such a way that the sums of the contributions of all genes will equal as closely as possible the actual measures of the character in the individuals of the population. Obviously there could be exact agreement in all cases only if dominance and epistatic relationships were completely lacking. ...[W]ith respect to such a character as fitness, it may safely be assumed that there are always important epistatic effects. Genes favorable in one combination, are, for example extremely likely to be unfavorable in another." (Wright, 1930)

Wright's protests notwithstanding, the assumption that the fitness effects of allele

substitutions are context-independent has become the orthodoxy in population genetics. This position is described on the first page of a popular introduction to population genetics as follows:

"Population geneticists have achieved remarkable success by choosing to ignore the complexities of real populations and focusing on the evolution of one or a few loci at a time. . . . The success of this approach, which has been seen in both theoretical and experimental investigations has been impressive, as I hope the reader will agree by the end of this book. The approach is not without its detractors. Years ago, Ernst Mayr mocked this approach as 'bean bag genetics.' In so doing, he echoed a view held by many of the pioneers of our field that natural selection acts on highly interactive coadapted genomes whose evolution cannot be understood by considering the evolution of a few loci in isolation from all others. Although genomes are certainly coadapted, there is precious little evidence that there are strong interactions between most polymorphic alleles in natural populations. The modern view, spurred on by the rush of DNA sequence data, is that we can profitably study loci in isolation." (Gillespie, 1998)

This "modern view" is currently challenged by a small, but outspoken community of critics, who consider the practice of theoretically analyzing isolated loci to be driven by convenience and scandalously naive; see, for example, the volume by Wolf et. al (Wolf et al., 2000). Consider the following passage from this compilation which accounts for Gillespie's observation that "there is precious little evidence that there are strong interactions between most polymorphic alleles in natural populations". Paraphrasing statements made by Frankel and Shork (Frankel and Schork, 1996), Templeton writes:

"The subjective assessment of Frankel and Shork [. . .] implies that epistasis is common, despite the numerous biases that exist against its detection. Frankel and Shork [. . .] point out that the primary reason why many complex traits are not reported to have epistasis is simply that many investigators use designs and/or analytical methods that exclude epistasis. The implicit assumption in these analyses is that all the biologically important associations are to be found at the single-locus level." (Templeton, 2000)

Later in the same paper, Templeton cautions that the pervasiveness of this "implicit assumption" amounts to a community-wide neglect of epistasis because of the mathematical and statistical inconvenience it poses.

"The dominance of this research paradigm has more to do with mathematical and statistical convenience than with biological reality. All that we know of biological systems—from the control of gene expression, to biochemical pathways, to developmental processes, to physiological regulation—indicates that interactions are the norm." (Templeton, 2000)

Meanwhile Rice has argued that the very notion of epistasis is questionable. Not because epistasis does not exist, but because it is ubiquitous.

" 'Epistasis,' like 'invertebrate,' is a term that really means 'everything else'. Traditionally defined as a situation in which the consequences of an allele substitution at one locus are a function of what allele is present at another locus

..., epistasis includes all possible ways that gene products can conspire to shape a phenotype, with the very unlikely exception of complete additivity. To name a phenomenon in this way has the curious effect of making it look like a special case, even if it is the most common situation."(Rice, 2000)

## 5.2 Comparing Fisher's Assumptions with Holland's

Proponents of the building block hypothesis typically regard each symbol in a genomic string as a separate gene (Holland, 1975; Goldberg, 1989; Mitchell, 1996). A building block can then be (incompletely) described as a small set of closely located genes amongst which credit cannot be apportioned, i.e. a set of genes with epistatic interactions. As building blocks are assumed to be abundant, proponents of the building block hypothesis feel justified in claiming that the building block hypothesis accommodates the existence of pervasive epistasis, and therefore rests on assumptions that are much weaker than Fisher's.

A less cozy picture however emerges if one defines a gene more in line with its definition in population genetics. The word "gene" was coined by the Danish geneticist Wilhelm Johansen in 1909, at a time when the particulate nature of inheritance could only be surmised by observing differences in the frequencies of some phenotypic trait in an population over multiple generations. Johansen used the word gene to stand for a *unit of inheritance* passed on from parent to child in an all-or-nothing fashion. The theories of the founders of population genetics—Fisher, Wright, Haldane—all make use of these fictitious units. Later, as the physical basis of inheritance started to become clear, molecular biologists began to identify a gene with any regulatory, transcribed, or other functional region of DNA. Unfortunately this use of the term "gene" is not consistent with its use in population genetics. In order to maintain consistency with the theoretical work of their field's founders, population geneticists typically identify a gene with a chromosomal extent that tends to be inherited in an all-or-nothing fashion, i.e. a chromosomal extent that is short enough that it tends not to be broken up by crossover (Dawkins, 1999, p28). This is not a strict definition, like say that of a triangle, but instead has a "fading-out" quality that is contingent upon the expected number of crossover points and the way they tend to be distributed over the genome[2].

If we now consider GAs with $n$-point crossover, where $n$ is small, then, in light of the above, any short schema with contiguous defining positions is a gene. Let us adopt this definition of a gene. Note then that any building block must be comprised of one or more genes with above average fitness. For example, if ***************10*1*** is a building block, then one or both of the genes ***************1001***, and ***************1011*** must have above average fitness. It then becomes straightforward to see how the building block hypothesis and the assumptions that undergird it can be described entirely in terms of genes.

We are now in a position to compare Fisher's assumptions with Holland's. Both Fisher and Holland assumed the existence of a large number of genes with higher than average fitness. Fisher essentially assumed that *any* collection of such genes intersects synergistically. Holland, assumed a) that synergistic intersections between small co-located collections of such genes is common, b) that antagonistic intersections between a collection of such genes is rare, and c) that this pattern applies hierarchically. To be sure these three assumptions about the distribution of fitness *are* weaker than Fisher's. However, given the extraordinary strength of Fisher's assumptions, that's not saying

---

[2]Dawkins uses the word *cistron* to refer to what molecular biologists call a gene (Dawkins, 1999, p28).

13

much.

Fisher and Holland also differ in the way they deal with the problem of sampling error. By assuming an *infinite* panmictic population, Fisher dispensed with the need for the average fitness values of the alleles at a locus to be detectably different. This is because in an infinite panmictic population, evolution can act on differences between the average fitness values of these alleles no matter how small the differences. In other words, by making a strong assumption about the size of the evolving population, Fisher was able to avoid making a strong assumption about the distribution of fitness. Genetic algorithms used in practice tend to have small populations—typically no more than 1000 individuals. Therefore an escape of this sort is clearly not available to genetic algorithm theorists

Nevertheless, to the best of our knowledge the issue of sampling error has not been addressed by proponents of the building block hypothesis. By insisting that the building block hypothesis explains the adaptive capacity of GAs with small populations, and by making no effort to address the issue of *detectability* that arises when one takes this position, these proponents are, in effect, making the assumption that each basic building block is comprised of at least one gene whose sampling fitness is likely to be so far above average that evolution will propagate this gene despite the inevitable sampling error that accompanies the evolution of small populations. The assumption that basic building blocks are abundant entails the strong assumption that such genes are also abundant.

## 6 The Problem with "The Loose Linkage Problem"

While proponents of the building block hypothesis seem to be comfortable with the strong assumption that low-order schemata with above average sampling fitness are abundant, they are less enthusiastic about assuming abundance when it is also stipulated that these schemata must be short. In other words, these proponents are uncomfortable assuming that the defining positions of basic building blocks are "tightly linked", i.e. close together.

Their wariness about this assumption can be traced back to Holland's contention in his seminal treatise (Holland, 1975) that the defining bits of building blocks may well be dispersed throughout the genome (i.e. loosely linked). Holland considered this to be a significant problem. To deal with it he introduced the *inversion* operator which reverses the order of the bits of a randomly chosen snippet of a genome while preserving the genome's semantics (Holland, 1975, p106). Holland asserted that over several generations inversion, in combination with crossover and selection, will tighten the linkage between the defining positions of schemata with above average fitness in an "intrinsically parallel fashion" (Holland, 1975, p109).

The inversion operator was not found to be useful in practice (Davis, 1991), and the study of inversion is no longer an active area of research. However, the loose linkage problem that inversion was supposed to solve took on a life of its own—it became a *de-facto* explanation for poor GA performance, and has captured the attention and creative energies of a sizeable section of the GA community. A large number of algorithms with explicit "linkage learning" schemes have been developed to deal with this perceived problem e.g. mGA and the fmGA (Goldberg et al., 1989, 1990, 1993; Kargupta, 1995), gemGA, (Kargupta, 1996; Goldberg, 2002), LLGA (Harik and Goldberg, 1997; Goldberg, 2002), ECGA (Harik, 1999), FDA (Mühlenbein and Mahnig, 1999), LFDA

(Mühlenbein and Mahnig, 2001), BOA (Pelikan et al., 1999; Goldberg, 2002), hBOA (Pelikan and Goldberg, 2001), SEAM (Watson, 2002, 2006).

What seems to have been overlooked in this flurry of activity is that the assumption of tight linkage between the defining positions of basic building blocks is just one of a number of strong assumptions undergirding the building block hypothesis. Consciously or otherwise, these other assumptions—the *abundance* of basic building blocks, and hierarchical synergism—have been embraced by the inventors of the algorithms listed above.

Goldberg calls such algorithms "competent genetic algorithms" (Goldberg, 2002). The implication, of course, is that the simple genetic algorithm, which lacks explicit linkage learning mechanisms, is *in*competent. Driving home this point in one of his books, Goldberg writes:

> "One mistaken idea that has led to controversy is the idea that simple GAs as originally designed [. . . ] or their minor variants achieve the kind of robustness sought in Holland's early writing. This text puts this canard to rest; even my first text went to great lengths to discuss the importance of linkage and the inadequacy of simple GAs in solving the linkage problem. Nonetheless, the field has proceeded using simple GAs as though they worked well. They do not." (Goldberg, 2002, p55)

Sadly, Goldberg does not consider the possibility that simple GAs and minor variants thereof continue to be used by GA practitioners because they *do* "work well", but *don't* work as described in the building block hypothesis.

An unfortunate consequence of the field's preoccupation with the loose linkage "problem" is the diversion of effort that might otherwise have been devoted to developing more viable explanations for the adaptive capacity of the simple genetic algorithm. Why, after all, would an engineer care to study the workings of an algorithm deemed to be the poor cousin of more "competent" algorithms?

## 7  Conclusion

What is a genetic algorithm? Is it

**(a)** An algorithmic model of sexually reproducing biological populations, which, despite its abstract nature, retains something of the remarkable capacity for adaptation possessed by such populations.

  or

**(b)** An algorithm that implements the process described in the building block hypothesis

Proponents of the building block hypothesis regard both (a) and (b) as valid descriptions of a genetic algorithm. However, of the two descriptions, their loyalty seems to reside with description (b). Holland, for example, writes "the very essence of good GA design is retention of diversity, furthering exploration, while exploiting building blocks already discovered" (Holland, 2000)[3]. Most researchers in the field however

---

[3]One has to wonder if the building block hypothesis can really be called a hypothesis if essential parts of it are incorporated into the *definition* of a GA.

readily agree with description (a), but need to be convinced that (b) is also a valid description. The theoretical basis of the argument by which BBH proponents attempt to convince us of the validity of description (b) has been sharply criticized. Unfortunately criticism of this sort has been brushed aside by BBH proponents who continue to insist that (b) is a valid description.

"What is it that people do when they are being innovative in a cross-fertilizing sense?", asks Goldberg. "Usually they are grasping at a notion—a set of good solution features—in one context, and a notion in another context and juxtaposing them, thereby speculating that the combination might be better than either notion taken individually" (Goldberg, 2002, p5). Goldberg calls the *de-facto* attribution to genetic algorithms of this purported process of human innovation the *fundamental intuition of genetic algorithms*, or the *innovation intuition*.

We find Goldberg's attempts to rationalize this attribution—for example, by quoting the French mathematician Jacques Hadamard: "Indeed, it is obvious that invention or discovery be it in mathematics or anywhere else, takes place by combining ideas" (Hadamard, 1954)—unconvincing.

No one can deny the aesthetic appeal of the building block hypothesis, least of all engineers, weaned (as we typically are) on the virtues and use of modularity. That genetic algorithms might be constructing solutions to problems much as engineers do— by identifying important modules, and composing them together—is a truly exciting prospect. Unfortunately, as discussed in this paper, some extraordinarily strong assumptions about the distribution of fitness have to hold true in order for this prospect to be realized.

The history of science abounds with conceptual entities that were highly popular for a while, but were later jettisoned because of the extraordinary strength of the assumptions one had to embrace. One such entity is *luminiferous aether*, which, for much of the nineteenth century, was assumed to be the medium through which light travels. Light, like all waves, bends around objects (diffraction), changes direction upon striking a reflective surface (reflection) or entering a new medium (refraction), can be split into components with different frequencies (dispersion), and displays interference patterns. Physicists in the nineteenth century assumed that the propagation of light, like the propagation of all waves known at the time, requires the mechanical disturbance of some physical medium. This medium was called luminiferous (i.e. light bearing) aether.

On the one hand the concept of a luminiferous aether proved very useful because it allowed physicists to account for phenomena like diffraction, which presented serious problems for a corpuscular theory of light; on the other hand this concept presented some serious difficulties of its own. For example, the aether had to be extremely rigid in order to support the high frequencies of light. At the same time, because it did not seem to have any observable effect on the orbits of the planets, it had to be devoid of mass and viscosity! Paradoxes like these occupied the minds of some of the finest physicists during the latter half of nineteenth century and into the early twentieth. Ultimately, the existence of the aether was obviated by the less presumptive theories of Maxwell (who cast light as a electromagnetic wave, *not* dependent on the mechanical properties of an aether), and Einstein (specifically his work on wave-particle duality and special relativity). The luminiferous aether, in other words, fell to Occam's razor.

Will building blocks go the way of luminiferous aether? Time will tell. At this point what we *can* say is that certain parallels between the two are unmistakeable. Both are

the result of extrapolation—of the mechanical basis of wave propagation in the second case, and of the purported process underlying human innovation in the first—and both require believers to embrace, whether consciously or otherwise, assumptions so strong, they seem almost magical.

## References

D. J. Cavicchio. *Adaptive Search Using Simulated Evolution*. PhD thesis, University Of Michigan, Ann Arbor, 1970.

B. Jack Copeland. Computation. In Luciano Floridi, editor, *The Blackwell Guide to the Philosopy of Computing and Information*, chapter 1, pages 3–17. Blackwell Publishing, 2004.

Lawrence Davis. *Handbook of genetic algorithms*. Van Nostrand Reinhold, New York, 1991.

Richard Dawkins. *The Selfish Gene*. Oxford University Press, 1999.

Kenneth De Jong. *An Analysis of the Behavior of a class of genetic adaptive systems*. PhD thesis, University of Michigan, Ann Arbor, 1975.

Kenneth De Jong. *Evolutionary Computation: A Unified Approach*. MIT Press, 2006.

RA Fisher. *The genetical theory of natural selection. Clarendon*. Oxford, 1930.

David B. Fogel. *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*. IEEE Press, NY, 2000.

Stephanie Forrest and Melanie Mitchell. Relative building-block fitness and the building-block hypothesis. In L. Darrell Whitley, editor, *Foundations of Genetic Algorithms 2*, pages 109–126, San Mateo, CA, 1993. Morgan Kaufmann.

W. N. Frankel and N. J. Schork. Who's afraid of epistasis. *Nature Genetics*, 14:371–373, 1996.

John H. Gillespie. *Population Genetics: A Concise Guide*. The John Hopkins University Press, 1998.

D. E. Goldberg, B. Korb, and K. Deb. Messy genetic algorithms: Motivation, analysis, and first results. *Complex Systems*, 3:493–530, 1989.

David E. Goldberg. *Genetic Algorithms in Search, Optimization & Machine Learning*. Addison-Wesley, Reading, MA, 1989.

David E. Goldberg. *The Design Of Innovation*. Kluwer Academic Publishers, 2002.

David E. Goldberg, Kalyanmoy Deb, and Bradley Korb. Messy genetic algorithms revisited: Studies in mixed size and scale. *Complex Systems*, 4(4):415–444, 1990.

David E. Goldberg, Kalyanmoy Deb, Hillol Kargupta, and Georges Harik. Rapid, accurate optimization of difficult problems using fast messy genetic algorithms. pages 56–64, 1993.

J. Hadamard. *The Psychology of Invention in the Mathematical Field*. Dover, NY, 1954.

G. Harik. Linkage learning via probabilistic modeling in the ecga, 1999.

Georges R. Harik and David E. Goldberg. Learning linkage. In Richard K. Belew and Michael D. Vose, editors, *Foundations of Genetic Algorithms 4*, pages 247–262, San Francisco, CA, 1997. Morgan Kaufmann.

John H. Holland. Genetic algorithms. *Scientific American*, 267, 1992.

John H. Holland. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. MIT Press, 1975.

John H. Holland. Building blocks, cohort genetic algorithms, and hyperplane-defined functions. *Evolutionary Computation*, 8(4):373–391, 2000.

Hollstien. *Artificial Genetic Adaptation in Computer Control Systems*. PhD thesis, University of Michigan, Ann Arbor, 1971.

Hillol Kargupta. The gene expression messy genetic algorithm. pages 814–819, 1996.

Hillol Kargupta. *SEARCH, Polynomial Complexity, and the Fast Messy Genetic Algorithm*. PhD thesis, University of Illinois at Urbana-Champaign, 1995.

T.S. Kuhn. *The Structure of Scientific Revolutions (: 1996)*.

Ernst Mayr. *The growth of biological thought*. Belknap Press of Harvard University Press, 2003.

Melanie Mitchell. *An Introduction to Genetic Algorithms*. The MIT Press, Cambridge, MA, 1996.

H. Mühlenbein and T. Mahnig. Evolutionary algorithms: From recombination to search distributions. In Leila Kallel, Bart Naudts, and Alex Rogers, editors, *Theoretical Aspects of Evolutionary Computing*, pages 135–173, Berlin, 2001. Springer.

Heinz Mühlenbein and Thilo Mahnig. The factorized distribution algorithm for additively decompressed functions. In *1999 Congress on Evolutionary Computation*, pages 752–759, Piscataway, NJ, 1999. IEEE Service Center.

A.E. Nix and M.D. Vose. Modeling genetic algorithms with Markov chains. *Annals of Mathematics and Artificial Intelligence*, 5(1):79–88, 1992.

S. Okasha. *Philosophy of Science: A Very Short Introduction*. Oxford University Press, 2002.

Karl Pearson and Alice Lee. On the laws of inheritance in man. 1. inheritance of physical characters. *Biometrika*, 2:357–462, 1903.

Martin Pelikan and David E. Goldberg. Hierarchical bayesian optimization algorithm = bayesian optimization algorithm + niching + local structures. In *Optimization by Building and Using Probabilistic Models (OBUPM) 2001*, pages 217–221, San Francisco, California, USA, 7 July 2001.

Martin Pelikan, David E. Goldberg, and Erick Cantú-Paz. BOA: The Bayesian optimization algorithm. In *Proceedings of the Genetic and Evolutionary Computation Conference GECCO-99*, volume I, pages 525–532. Morgan Kaufmann Publishers, San Fransisco, CA, 13-17 July 1999.

Karl Popper. *Conjectures and Refutations*. Routledge, 2007a.

Karl Popper. *The Logic Of Scientific Discovery*. Routledge, 2007b.

William B. Provine. *The Origins of Theoretical Population Genetics*. The University of Chicago Press, Chicago, 2001.

A. Prügel-Bennett and J.L. Shapiro. Analysis of genetic algorithms using statistical mechanics. *Physical Review Letters*, 72(9):1305–1309, 1994.

Magus L. Rattray. *Modelling the Dynamics of a Genetic Algorithm Using Statistic Mechanics*. PhD thesis, Computer Science Department, University of Manchester, 1996.

C.R. Reeves and J.E. Rowe. *Genetic Algorithms: Principles and Perspectives: a Guide to GA Theory*. Kluwer Academic Publishers, 2003.

Sean H. Rice. The evolution of developmental interactions. In Wolf et al. (2000).

J. L. Shapiro. Statistical mechanics theory of genetic algorithms. In Leila Kallel, Bart Naudts, and Alex Rogers, editors, *Theoretical Aspects of Evolutionary Computing*, pages 87–108, Berlin, 2001. Springer.

H. Simon. *The Sciences of the Artificial*. The M.I.T. Press, 1969.

Alan R. Templeton. Epistasis and complex traits. In Jason B. Wolf, Edmund D. Brodie, and Michael J. Wade, editors, *Epistasis and the Evolutionary Process*. Oxford University Press, 2000.

A. Tversky and D. Kahneman. Judgment under Uncertainty: Heuristics and Biases. *Science*, 185(4157):1124–1131, 1974.

Michael D. Vose. Modeling simple genetic algorithms. In L. Darrell Whitley, editor, *Foundations of Genetic Algorithms 2*, pages 63–73, San Mateo, 1993. Morgan Kaufmann.

Michael D. Vose. *The simple genetic algorithm: foundations and theory*. MIT Press, 1999.

Michael D. Vose and Gunar E. Liepins. Punctuated equlibria in genetic search. *Complex Systems*, 5(1):31–44, February 1991.

Richard A. Watson. *Compositional Evolution: Interdisciplinary Investigations in Evolvability, Modularity, and Symbiosis*. PhD thesis, April 01 2002.

Richard A. Watson. *Compositional Evolution: The Impact of Sex, Symbiosis and Modularity on the Gradualist Framework of Evolution*. The MIT Press, 2006.

Jason B. Wolf, Edmund D. Brodie III, and Michael J. Wade, editors. *Epistasis And the Evolutionary Process*. Oxford University Press, 2000.

Alden H. Wright, Michael D. Vose, and Jonathan E. Rowe. Implicit parallelism. In *GECCO*, 2003.

S. Wright. Review of The Genetical Theory of Natural Selection, by RA Fisher. *Journal of Heredity*, 21:80–87, 1930.