

LIFE'S CONSERVATION LAW

Why Darwinian Evolution Cannot Create Biological Information

William A. Dembski and Robert J. Marks II

Evolutionary Informatics Lab
www.evoinfo.org

- 1 The Creation of Information
- 2 Biology's Information Problem
- 3 The Darwinian Solution
- 4 Computational vs. Biological Evolution
- 5 Active Information
- 6 Three Conservation of Information Theorems
- 7 The Law of Conservation of Information
- 8 Applying LCI to Biology
- 9 Conclusion: "A Plan for Experimental Verification"

ABSTRACT: Laws of nature are universal in scope, hold with unfailing regularity, and receive support from a wide array of facts and observations. The Law of Conservation of Information (LCI) is such a law. LCI characterizes the information costs that searches incur in outperforming blind search. Searches that operate by Darwinian selection, for instance, often significantly outperform blind search. But when they do, it is because they exploit information supplied by a fitness function—information that is unavailable to blind search. Searches that have a greater probability of success than blind search do not just magically materialize. They form by some process. According to LCI, any such search-forming process must build into the search at least as much information as the search displays in raising the probability of success. More formally, LCI states that raising the probability of success of a search by a factor of q/p (> 1) incurs an information cost of at least $\log(q/p)$. LCI shows that information is a commodity that, like money, obeys strict accounting principles. This paper proves three conservation of information theorems: a function-theoretic, a measure-theoretic, and a fitness-theoretic version. These are representative of conservation of information theorems in general. Such theorems provide the theoretical underpinnings for the Law of Conservation of Information. Though not denying Darwinian evolution or even limiting its role in the history of life, the Law of Conservation of Information shows that Darwinian evolution is inherently teleological. Moreover, it shows that this teleology can be measured in precise information-theoretic terms.

1 The Creation of Information

Any act of intelligence requires searching a space of possibilities to create information. The preceding sentence illustrates this very point. In formulating it, we searched a space of letter sequences. Most such sequences are gibberish (unpronounceable arrangements of random letters). Of those that are meaningful, most are not in English. Of those that are in English, most have nothing to do with information theory or intelligent design. Only by successively eliminating vast swatches of this search space did we succeed in locating the first sentence of this paper, thereby creating a unique sentence in the history of the English language. Each such reduction of possibilities constitutes an information-generating act. Together, these reductions constitute a search that identifies one possibility (the first sentence of this paper) to the exclusion of others.

Information identifies possibilities to the exclusion of others. Unless possibilities are excluded, no information can be conveyed. To say “it’s raining or it’s not raining” is uninformative because that statement excludes no possibilities. On the other hand, to say “it’s raining” excludes the possibility “it’s not raining” and therefore conveys information. Tautologies, because they are true regardless, can convey no information. We don’t need to be informed of them because we can figure them out on our own.

Information presupposes multiple live possibilities. Robert Stalnaker puts it this way: “Content requires contingency. To learn something, to acquire information, is to rule out possibilities. To understand the information conveyed in a communication is to know what possibilities would be excluded by its truth.”¹ Fred Dretske elaborates: “Information theory identifies the amount of information associated with, or generated by, the occurrence of an event (or the realization of a state of affairs) with the reduction in uncertainty, the elimination of possibilities, represented by that event or state of affairs.”²

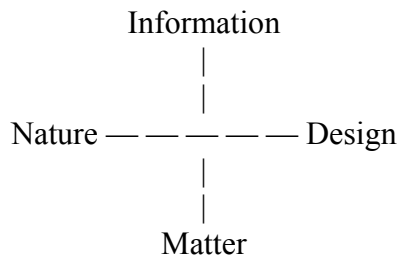
According Douglas Robertson, the defining feature of intelligent agents is their ability, as an act of free will, to create information.³ G. K. Chesterton elaborated on this insight: “Every act of will is an act of self-limitation. To desire action is to desire limitation. In that sense every act is an act of self-sacrifice. When you choose anything, you reject everything else.... Every act is an irrevocable selection and exclusion. Just as when you marry one woman you give up all the others, so when you take one course of action you give up all the other courses.”⁴ Intelligence creates information.

But is intelligence the only causal power capable of creating information? Darwin’s main claim to fame is that he is supposed to have provided a mechanism that could create information without the need for intelligence. Interestingly, he referred to this mechanism as “natural selection.” Selection, as understood before Darwin, had been an activity confined to intelligent agents. Darwin’s great coup was to attribute the power of selection to nature—hence “natural selection.”

Nature, as conceived by Darwin and his followers, acts without purpose—it is nonteleological and therefore unintelligent. As evolutionary geneticist Jerry Coyne puts it in opposing intelligent design, “If we’re to defend evolutionary biology, we must defend it as a science: a *nonteleological* theory in which the panoply of life results from the action of natural selection and genetic drift acting on random mutations.”⁵ But why do Coyne and fellow Darwinists insist that evolutionary biology, to count as science, must be nonteleological.⁶ Where did that rule come from? The wedding of teleology with the natural sciences is itself a well-established science—it’s called engineering. Intelligent design, properly conceived, belongs to the engineering sciences.

But to return to the point at hand, does nature really possess the power to select and thereby

create information? To answer this question, we need to turn to the relation between matter and information. The matter-information distinction is old and was understood by the ancient Greeks. For them, there was matter, passive or inert stuff waiting to be arranged; and there was information, active stuff that did the arranging.⁷ This distinction provides a useful way of carving up experience and making sense of the world. Left here, it is uncontroversial. Nonetheless, it becomes controversial once we add another dimension to it, that of nature and design:



Designing intelligences are not the only agents capable of putting information into matter. Nature, too, has this capacity. Consider the difference between raw pieces of wood and an acorn. Raw pieces of wood do not have the power to assemble themselves into a ship. For raw pieces of wood to form a ship requires a designer to draw up a blueprint and then take the pieces of wood and, in line with the blueprint, fashion them into a ship. But where is the designer that causes an acorn to develop into a full-grown oak tree? The acorn possesses the power to transform itself into an oak tree.

Nature and design therefore represent two different ways of producing information. Nature produces information internally. The acorn assumes the form it does through powers it contains within itself—the acorn is a seed programmed to produce an oak tree. Accordingly, the acorn does not create its information from scratch. Rather, it expresses already existing information, which it derived from previous generations of oak trees. By contrast, a ship assumes the form it does through powers external to it—a designing intelligence needs to intentionally arrange the pieces of wood to form a ship.

Not only did the ancient Greeks know about the distinction between information and matter, but they also knew about the distinction between design and nature. Thus Aristotle characterized design as consisting of capacities external to objects that are needed to bring about their form. On the other hand, he saw nature as consisting of capacities internal to objects that enable them to transform themselves without outside help. Thus, in Book XII of the *Metaphysics*, Aristotle wrote, “[Design] is a principle of movement in something other than the thing moved; nature is a principle in the thing itself.”⁸ In Book II of the *Physics* Aristotle referred to design as completing “what nature cannot bring to a finish.”⁹

The Greek word here translated “design” is *techne*, from which we get our word technology. In translations of Aristotle’s work, the English word most commonly used to translate *techne* is “art” (in the sense of “artifact”). Design, art, and *techne* are synonyms. The essential idea behind these terms is that information is conferred on an object from outside it and that the materials that make up the object lack the power to organize themselves into it. Thus, apart from that outside information, the object cannot assume the form it does. For instance, raw pieces of wood do not by themselves have the power to form a ship.

This contrasts with nature, which does have the power within itself to express information. Thus, in Book II of the *Physics*, Aristotle wrote, “If the ship-building art were in the wood, it would produce the same results by nature.”¹⁰ In other words, if raw pieces of wood had the

capacity to form ships, we would say that ships come about by nature. The Greek word here translated “nature” is *phusis*, from which we get our word physics. The Indo-European root meaning behind *phusis* is growth and development. Nature produces information not by imposing it from outside but by growing or developing informationally rich structures from the capacities inherent in a thing. Consider again the acorn. Unlike wood that needs to be fashioned by a designer to form a ship, acorns produce oak trees naturally—the acorn simply needs a suitable environment in which to grow.

Aristotle claimed that the art of ship-building is not in the wood that makes up a ship. In same vein, the art of sentence-making is not in the letters of the alphabet out of which sentences are constructed. Likewise, the art of statue-making is not in the stone out of which statues are made. Each of these cases requires a designer. So too, the theory of intelligent design contends that the art of life-making is not in the physical stuff that constitutes life. Instead, life-making requires a designer.

The central issue in the scientific debate over intelligent design and biological evolution can therefore be stated as follows: Is nature complete in the sense of possessing all the resources it needs to bring about the information-rich biological structures we see around us or does nature also require some contribution of design to bring about those structures? Darwinian naturalism argues that nature is able to create all its own information and is therefore complete. Intelligent design, by contrast, argues that nature is merely able to reexpress existing information and is therefore incomplete.

To argue for nature’s incompleteness seems to presuppose a fundamental distinction between intelligent and natural causes. Can’t intelligent causes be fully natural? As the scientific community understands them, natural causes operate according to deterministic and nondeterministic laws and can be characterized in terms of chance, necessity, and their combination.¹¹ To be sure, if one is more liberal about what one means by natural causes and includes among them end-directed (teleological) processes that are not reducible to chance and necessity (as Aristotle and the ancient Stoics did by endowing nature with immanent teleology), then our claim that natural causes are incomplete dissolves.

But for many in the scientific community, natural causes are at heart nonteleological and therefore unintelligent. Natural causes, when suitably coordinated, may exhibit intelligence. Thus animals might be viewed as purely natural objects that act as intelligent causes to achieve ends. But the teleology they exhibit is, from a naturalistic perspective, the result of a long and blind evolutionary process whose nuts-and-bolts causal processes are nonteleological. Given naturalism, natural causes bring about intelligent causes but are not themselves intelligent. On this view, intelligent causes are always reducible to nonteleological natural causes, ultimately to the motions and interactions of particles governed by forces of attraction and repulsion.

The challenge of intelligent design, and of this paper in particular, is to show that when natural systems exhibit intelligence by producing information, they have in fact not created it from scratch but merely shuffled around existing information. Nature is a matrix for expressing already existent information. But the ultimate source of that information resides in an intelligence not reducible to nature. The Law of Conservation of Information, which we explain and justify in this paper, demonstrates that this is the case. Though not denying Darwinian evolution or even limiting its role as an immediate efficient cause in the history of life, this law shows that Darwinian evolution is deeply teleological. Moreover, it shows that the teleology inherent in Darwinian evolution is scientifically ascertainable—it is not merely an article of faith.

2 Biology's Information Problem

Caltech president and Nobel Prize-winning biologist David Baltimore, in describing the significance of the Human Genome Project, stated, "Modern biology is a science of information."¹² Biologists now recognize the crucial importance of information to understanding life and, especially, its origin. Origin-of-life researcher and fellow Nobel Prize winner Manfred Eigen equated the problem of life's origin with uncovering "the origin of information."¹³ Biologists John Maynard Smith and Eörs Szathmáry have explicitly placed information at the center of developmental and evolutionary biology: "A central idea in contemporary biology is that of information. Developmental biology can be seen as the study of how information in the genome is translated into adult structure, and evolutionary biology of how the information came to be there in the first place."¹⁴

Given the importance of information to biology, the obvious question is, How does biological information arise? For matter to be alive it must be suitably structured. A living organism is not a mere lump of matter. Life is special, and what makes it special is how its matter is arranged into very specific forms. In other words, what makes life special is information. Where did the information necessary for life come from?

The emergence of life constitutes a revolution in the history of matter. A vast gulf separates the organic from the inorganic world, and that gulf is properly characterized in terms of information. The matter that makes up the bricks of your house and the matter that makes up your body is essentially the same. Nevertheless, the arrangement of that matter—the information—differs vastly in these two cases. Biology's information problem is therefore to determine whether (and, if so, how) purely natural forces are able to bridge the gulf between the organic and inorganic worlds as well as the gulfs between different levels of complexity within the organic world. Conversely, biology's information problem is to determine whether (and, if so, how) design is needed to complement purely natural forces in the origin and subsequent development of life.

But how can we determine whether nature has what it takes to produce life? The sheer commonness of life on earth tends to lull us into complacency. We look around and see life everywhere. But there was a time when the earth contained no multi-celled organisms like us. Before that, the earth contained no life at all, not even single-celled forms. And earlier still, there was no earth at all, no sun or moon or sister planets. Indeed, if physicists are right, there was a time when there were no stars or galaxies but only elementary particles like quarks densely packed at incredibly hot temperatures. That would coincide with the moment just after the Big Bang.

Suppose we go back to that moment. Given the history of the universe since then, we could say—in retrospect—that all the possibilities for complex living forms like us were in some sense present at that earlier moment in time (much as many possible statues are in some sense present in a block of marble). From that early state of the universe, galaxies and stars eventually formed, then planet earth, then single-celled life forms, and finally life forms as complicated as us. But that still doesn't tell us how we got here or whether nature had sufficient creative power to produce us apart from design. Nature provides a necessary condition for our existence. The question is whether it also provides a sufficient condition.

As philosopher Holmes Rolston points out, humans are not invisibly present in primitive single-celled organisms in the same way that an oak tree is secretly present in an acorn. The oak tree unfolds in a lawlike or programmatic way from an acorn. But the same cannot be said for the grand evolutionary story that places single-celled organisms at one end and humans at the other ("monad to man" evolution). There's no sense in which human beings or any other multi-celled

organisms are latent in single-celled organisms, much less in nonliving chemicals. For Rolston, the claim that life is somehow already present or lurking in nonliving chemicals or that complex biological systems are already present or lurking in simple biological systems is “an act of speculative faith.”¹⁵

Speculative faith is not science. It is unscientific simply to assert that nature possesses the causal powers capable of producing living forms. Rather, science has the job of demonstrating whether nature actually does possess such causal powers. Moreover, we do not have the luxury, like Aristotle and many ancient philosophers, of thinking that life and the universe has always been here. Advances in astrophysics and geology show that the earth has not been here forever and that the early earth was a tempestuous cauldron that rendered it uninhabitable. Yet somehow, from fairly simple inorganic compounds on the early lifeless earth, life forms requiring the precisely coordinated activity of information-rich biomacromolecules emerged. How did that happen? How could it have happened?

3 The Darwinian Solution

Karl Marx is said to have joked that the twenty-six letters of the alphabet were his soldiers for conquering the world.¹⁶ Yet to be successful, these soldiers required Marx’s supervision. Left to themselves, the letters of the alphabet are inert. Mere alphabetic characters do not have the power to transform themselves into Marx’s *Das Kapital*. Marx plus the letters of the alphabet, however, do have that power. But that raises the question how Marx himself came to be.

From a Darwinian perspective, any designing intelligence, such as Marx, results from a long and blind evolutionary process. Nature, without the need of any intelligence, starts off from a lifeless earth and, over the course of natural history, produces life forms that eventually evolve into human beings like Karl Marx, who then write economic treatises like *Das Kapital*. Within Darwinian naturalism, only natural forces, principally natural selection acting on random variations, control the evolutionary process. Designing intelligences may evolve out of that process but play no role in guiding or controlling it.

Theistic evolutionists attempt to make room for God within this Darwinian scheme by claiming that God created the universe so that Darwinian processes would produce living forms. Accordingly, God brings about the natural forces that produce living things but makes sure not to interfere with those forces once they are in operation. Though logically possible, theistic evolution offers no compelling reason for thinking that nature is a divine creation. As Francisco Ayala put it, “It was Darwin’s greatest accomplishment to show that the directive organization of living beings can be explained as the result of a natural process, natural selection, without any need to resort to a Creator or other external agent.”¹⁷ Darwinian evolution, far from requiring a creator God, is perfectly happy with a natural world that is entirely self-sufficient.

Theistic evolutionists think Darwin got nature right and then adapt their theology to suit Darwinian science. Proponents of intelligent design, by contrast, ask the logically prior question of whether Darwin did in fact get nature right. Indeed, why should we think that natural forces, apart from intelligent guidance, have the power to create biological information? Darwin attempted to resolve this question by providing a naturalistic mechanism (natural selection acting on random variations) that could effectively account for the production of biological information.

Some of Darwin’s followers consider this mechanism so effective that it makes unnecessary any intelligence behind biology. Thus Richard Dawkins even went so far as to state, “Darwin made it possible to be an intellectually fulfilled atheist.”¹⁸ Others, like Francis Collins, think that an intelligence (for him, the Christian God) set up the conditions that make it possible for the

Darwinian mechanism to bring about biological information. This view, according to Collins, “makes it possible for the scientist-believer to be intellectually fulfilled and spiritually alive.”¹⁹ Yet for Collins and other theistic evolutionists, God’s hand is nowhere evident in the evolutionary process. Atheistic and theistic evolution join hands in proclaiming that purposeful design in biology is *scientifically undetectable*.²⁰

How does the Darwinian mechanism replace intelligence in its account of biological information? Richard Dawkins uses an evolutionary algorithm to illustrate how the Darwinian mechanism is supposed to create information apart from intelligence.²¹ For convenience, we refer to this example as Dawkins’s WEASEL. He starts with the following target sequence:

METHINKS•IT•IS•LIKE•A•WEASEL

(he considers only capital Roman letters and spaces, spaces represented here by bullets—thus 27 possibilities at each location in a symbol string 28 characters in length).

If we tried to attain this target sequence by pure chance (for example, by randomly shaking out scrabble pieces consisting solely of letters and spaces), the probability of getting it on a given try would be around 1 in 10^{40} , and, correspondingly, it would take on average about 10^{40} tries to stand a better than even chance of getting it.²² Thus, if we depended on pure chance to attain this target sequence, we would in all likelihood be unsuccessful. To generate the linguistic information in Dawkins’s target sequence, pure chance is simply not up to the task.

Now the beauty of the Darwinian mechanism (it is a deceptive beauty) is that it is supposed to circumvent the limitations of pure chance. Thus, in place of pure chance, Dawkins considers the following evolutionary algorithm: (i) Start out with a randomly selected sequence of 28 capital Roman letters and spaces, e.g.,

WDL•MNL T•DTJBKWIRZREZLMQCO•P

(ii) randomly alter individual letters and spaces in the current sequence; (iii) keep those newly formed sequences that match more letters in the target sequence, discarding the rest. This algorithm exemplifies the Darwinian mechanism: step (ii) corresponds to the random variations that supply the raw potential for the evolutionary process; step (iii) corresponds to the selection of those variants that are better fitted to their environment (in this case, those letter sequences that match the target sequence more closely).

In very short order this algorithm converges to Dawkins’s target sequence. In *The Blind Watchmaker*, he summarizes a particular run of the algorithm that in a mere 43 steps converged to the target sequence:²³

```
(1)  WDL•MNL T•DTJBKWIRZREZLMQCO•P
(2)  WDLTMNL T•DTJB SWIRZREZLMQCO•P
      . . .
(10) MDLDMNLS•ITJISWHRZREZ•MECS•P
      . . .
(20) MELDINLS•IT•ISWPRKE•Z•WECSEL
      . . .
(30) METHINGS•IT•ISWLIKE•B•WECSEL
      . . .
(40) METHINKS•IT•IS•LIKE•I•WEASEL
      . . .
(43) METHINKS•IT•IS•LIKE•A•WEASEL
```

Thus, in place of 10^{40} tries on average for pure chance to produce the target sequence, by employing the Darwinian mechanism it now takes on average less than 100 tries to produce it. In short, a search effectively impossible for pure chance becomes eminently feasible for the Darwinian mechanism.

So does Dawkins's evolutionary algorithm demonstrate the power of the Darwinian mechanism to create biological information? No. Clearly, the algorithm was stacked to produce the outcome Dawkins was after. Indeed, because the algorithm was constantly gauging the degree of difference between the current sequence from the target sequence, the very thing that the algorithm was supposed to create (i.e., the target sequence METHINKS•IT•IS•LIKE•A•WEASEL) was in fact smuggled into the algorithm from the start. The Darwinian mechanism, if it is to possess the power to create biological information, cannot merely veil and then unveil existing information. Rather, it must create novel information from scratch. Clearly, Dawkins's algorithm does nothing of the sort.

Ironically, though Dawkins uses a targeted search to illustrate the power of the Darwinian mechanism, he denies that this mechanism, as it operates in biological evolution (and thus outside a computer simulation), constitutes a targeted search. Thus, after giving his METHINKS•IT•IS•LIKE•A•WEASEL illustration, he immediately adds: "Life isn't like that. Evolution has no long-term goal. There is no long-distant target, no final perfection to serve as a criterion for selection."²⁴ Dawkins here fails to distinguish two equally valid and relevant ways of understanding targets: (i) targets as humanly constructed patterns that we arbitrarily impose on things in light of our needs and interests and (ii) targets as patterns that exist independently of us and therefore regardless of our needs and interests. In other words, targets can be extrinsic (i.e., imposed on things from outside) or intrinsic (i.e., inherent in things as such).

In the field of evolutionary computing (to which Dawkins's METHINKS•IT•IS•LIKE•A•WEASEL example belongs), targets are given extrinsically by programmers who attempt to solve problems of their choice and preference. Yet in biology, living forms have come about without our choice or preference. No human has imposed biological targets on nature. But the fact that things can be alive and functional in only certain ways and not in others indicates that nature sets her own targets. The targets of biology, we might say, are "natural kinds" (to borrow a term from philosophy). There are only so many ways that matter can be configured to be alive and, once alive, only so many ways it can be configured to serve different biological functions. Most of the ways open to evolution (chemical as well as biological evolution) are dead ends. Evolution may therefore be characterized as the search for alternative "live ends." In other words, viability and functionality, by facilitating survival and reproduction, set the targets of evolutionary biology. Evolution, despite Dawkins's denials, is therefore a targeted search after all.

4 Computational vs. Biological Evolution

In the known physical universe, the number of elements that can be sampled (or queried) from a search space is always strictly limited. At the time of this writing, the fastest computer is the \$133-million IBM Roadrunner supercomputer at Los Alamos that operates at 1.059 petaflops (i.e., just over a thousand trillion floating point operations per second).²⁵ If we imagine each floating point operation as able to take a sample of size 1, then this computer, even when run over the duration of the physical universe (circa 12 billion years), would be able to sample at most $m = 10^{34}$ elements from the search space. It is estimated that the total number of organisms, both single-celled and multi-celled, that have existed on the earth over its duration (circa 4.5

billion years) is $m = 10^{40}$.²⁶ Thus it would take a million Roadrunner supercomputers running the duration of the universe to sample as many “life events” as have occurred on the earth.

Throughout this paper, we treat m as the upper limit on the number of elements that a given search can sample or query. Is there an upper limit to such upper limits? From examining the computational capacity of the universe at large, quantum computational theorist Seth Lloyd has shown that 10^{120} is the maximal number of bit operations that the known, observable universe could have performed throughout its entire multi-billion year history.²⁷ Thus, as the upper limit on the number of elements that a given search can sample, m cannot exceed 10^{120} , which sets an absolute limit on the sample size of any search.

Most search spaces that come up in the formation of biological complexity are far too large to be searched exhaustively. Take the search for a very modest protein, one that is, say, 100 amino acids in length (most proteins are several hundreds of amino acids in length). The space of all possible protein sequences that are 100 amino acids in length has size 20^{100} , or approximately 1.27×10^{130} , which exceeds Lloyd’s limit. For this space, finding a particular protein via blind search corresponds to a 1 in 10^{130} improbability. Exhaustively or blindly searching a space this size to find a target this small is utterly beyond not only present computational capacities but also the computational capacities of the universe as we know it.

Biochemist Robert Sauer has used a technique known as *cassette mutagenesis* to determine how much variation proteins can tolerate in their amino acids without disrupting function. His results show that taking this variation into account raises the probability of forming a 100-subunit functional protein to 1 in 10^{65} . But given 10^{65} atoms in our galaxy, this probability is still vanishingly small. Add to this that most proteins are not 100 but 250 to 300 amino acids in length and also that most proteins exist and operate in complexes requiring multiple proteins, and any prospect for blind search effectively exploring biological configuration space disappears.²⁸

Fortunately for evolutionary theory, Darwinian processes operating by natural selection are available to take the place of blind search and, so the theory contends, able to overcome its limitations. Darwinian search is thus supposed to counteract the vast improbabilities that at first blush seem to undermine the formation of biological complexity. Yet evolution, even when conceived as a Darwinian search, seems a terribly inefficient search strategy. All the significant innovations in biological form and complexity attributable to evolution are supposed to have taken thousands or even millions of years. Direct experimental verification of the ability of biological evolution to produce large-scale organismal change thus seems effectively impossible.

To accelerate the poky pace of biological evolution, early computer scientists recommended replacing the test-tube with the computer. Pioneers of evolutionary computing in the 1960s proposed that computer simulations could overcome the difficulty of demonstrating Darwinian evolution in the biology lab. “The Darwinian idea that evolution takes place by random hereditary changes and selection,” wrote Nils Barricelli in 1962, “has from the beginning been handicapped by the fact that no proper test has been found to decide whether such evolution was possible and how it would develop under controlled conditions.”²⁹

Whereas biological evolution occurred in deep time and therefore could not be observed, computers could model evolutionary processes in real time and thus render their behavior observable and open to experimental control. As J. L. Crosby put it back in the mid 1960s, “In general, it is usually impossible or impracticable to test hypotheses about evolution in a particular species by the deliberate setting up of controlled experiments with living organisms of that species. We can attempt to partially get around this difficulty by constructing [computer] models representing the evolutionary system we wish to study, and use these to test at least the theoretical validity of our ideas.”³⁰ Or, as Heinz Pagels summarized the matter two decades later,

“The only way to see evolution in action is to make computer models” because “in real time these changes take aeons, and experiment is impossible.”³¹

In the last two decades, however, confidence that computational evolution elucidates biological evolution has waned. Why is that? The short answer is that programmers can cook computer simulations to achieve any evolutionary result they want. Take intelligent-design critic Robert Pennock’s work on the computer program AVIDA. AVIDA, written by Christoph Adami, uses selection to evolve certain types of logic functions, which are viewed as virtual organisms. From the performance of this program, Pennock infers that evolutionary processes operating in nature can produce complex biological functions.³² Yet other computer programs, such as MESA, suggest that natural selection will have difficulty evolving features that need to form simultaneously for selective advantage. From MESA, Pennock might just as well have inferred that certain types of biological complexity (such as Michael Behe’s irreducibly complex molecular machines) may be unevolvable by Darwinian means.³³ So which program gives the better insight into biological evolution, AVIDA, which seems to confirm it, or MESA, which seems to disconfirm it?

It’s in large measure because computer programs can be manipulated to prove any evolutionary result one wants that ICAM was started. ICAM stands for the Institute for Complex Adaptive Matter.³⁴ Its mission is to understand how real-world material systems (as opposed to silicon-world virtual systems) become complex and adaptive. Talk to most working biologists, and they will tell you that computer simulations do not shed much light on actual biological evolution. Richard Lenski, Pennock’s collaborator on AVIDA, can appreciate this point.

Lenski is best known not for his work on computer programs that simulate biological evolution, but rather for his work as a straight-up biologist trying to evolve populations of bacteria in the lab. For many years, Lenski has cultured bacteria and placed them under selection pressure. He ran one experiment for 20,000 generations (if we think of the average generation time for humans as 20 years, then his experiment on bacteria corresponds to 400,000 years of human evolution, which is significant even on evolutionary time scales).³⁵ What did Lenski find in his experiments with these bacteria (i.e., with real, as opposed to virtual, organisms)? Did he find that his bacteria evolved novel irreducibly complex molecular machines of the sort that Michael Behe regards as indicating intelligent design?³⁶ Not at all. Lenski observed some small scale changes, but nothing remarkable. And yet, when Lenski turned to computer simulations, he found that virtual organisms are much easier to evolve than real ones, requiring only a few hundred generations to form novel complex structures.³⁷

Our ability to manipulate computer simulations of evolution has bred skepticism of the whole enterprise. Back around 1990, when artificial life was the most widely discussed form of computational evolution, John Maynard Smith called it “fact-free science.”³⁸ For David Berlinski, that skepticism has since turned to cynicism:

Computer simulations of Darwinian evolution fail when they are honest and succeed only when they are not. Thomas Ray has for years been conducting computer experiments in an artificial environment that he designated Tierra. Within this world, a shifting population of computer organisms meet, mate, mutate, and reproduce. Sandra Blakeslee, writing for the *New York Times*, reported the results under the headline “Computer ‘Life Form’ Mutates in an Evolution Experiment: Natural Selection is Found at Work in a Digital World.” Natural selection found at work? I suppose so, for as Blakeslee observes with solemn incomprehension, “the creatures mutated but showed only modest increases in complexity.” Which is to say, they showed nothing of interest at all. This *is* natural selection at work, but it is hardly work that has worked to intended effect.³⁹

Berlinski raises here an important question: what does it mean for a computer simulation of Darwinian evolution to succeed? For proponents of Darwinian evolution, success means that the simulations produce the types of structures and exhibit the types of capacities that biological evolution is supposed to have displayed in the course of natural history. Christoph Adami's AVIDA, Thomas Ray's Tierra, and, Thomas Schneider's *ev* (Schneider's program attempts to model the evolution of nucleotide binding sites) are thus supposed to constitute successes.⁴⁰ Indeed, as proponents of Darwinian evolution, Adami, Ray, and Schneider have a stake in seeing their simulations confirm standard evolutionary theory.

But critics of Darwinian evolution also write computer simulations. We've mentioned MESA (by William Dembski, John Bracht, and Micah Sparacio), but there are also others, for instance, Mendel's Accountant (by John Sanford and John Baumgardner) as well as the tongue-in-cheek MutationWorks (posted anonymously online).⁴¹ For such programs, success means showing that the simulation, despite efforts by the programmers to faithfully model biological evolution, produces no novel information or, worse yet, degrades existing information (as with Mendel's Accountant). Proponents of Darwinian evolution dismiss such simulations, claiming that their failure to evolve biologically relevant information merely reflects a failure of the programs to capture biological reality. Of course, critics of Darwinian evolution turn this criticism around, charging that only by cooking the books do Darwinists get their programs to produce biologically relevant information. Hence Berlinski's claim that these programs fail when they are honest.

Although it may seem as though we have reached an impasse, there is a way forward in this debate. Both sides have accepted a common presupposition, namely, that it is possible to model Darwinian evolutionary processes mathematically. Indeed, it better be possible to model evolution mathematically if it is to constitute an exact science.⁴² Yet the common presupposition that Darwinian processes can be modeled mathematically raises a logically prior question. We have considered whether particular mathematical models of Darwinian evolution can produce the sort of information we find in biological systems. Some appear capable of producing it, others not. The resulting debate centers on whether these models were tendentiously manipulated to achieve a desired result. But the logically prior question is whether, and in what sense, mathematical models of Darwinian processes allow for the production of biologically relevant information at all. The remainder of this paper takes up this question.

5 Active Information

Kenneth Miller, in reviewing Thomas Schneider's work on the computer simulation *ev*, attempts to account for the apparent increase in information that results from that algorithm. "What's needed to drive this increase?" he asks. "Just three things: selection, replication, and mutation." He continues,

Where's the new information coming from? Perhaps the investigator is sneaking it into the system at the start? No chance of that, since the starting sequences are completely randomized. Maybe there's hidden information in the program itself? Not likely. Schneider has made the source code of his program open for inspection, and there isn't even a hint of such nonsense. Did Schneider rig the parameters of the program to get the results he wanted? Not at all. In fact, changing the program in just about any way still results in an increase in measurable information, so long as we keep those three elements—selection, replication, and mutation—intact. Where the information "comes from" is, in fact, from the selective process itself.⁴³

This passage is remarkable for conceding what it seems to deny. The details of Schneider's simulation are not important here. We have argued elsewhere that *ev* is not nearly as free of investigator interference as Miller (or Schneider) makes out.⁴⁴ But let's grant, for the sake of argument, that Miller is right in denying investigator interference with the operation of the program. His claim that the information comes from the selective process is then correct but, in context, misleading. Miller suggests that *ev*, and evolution in general, outputs more information than it inputs. In fact, selective processes input as much information from the start as they output at the end. In Schneider's *ev*, for instance, the selective process inputted prior information in the form of a precisely specified error-counting function that served as a fitness measure.⁴⁵ Thus, instead of producing information in the sense of generating it from scratch, evolutionary processes produce it in the much weaker sense of merely shuffling around preexisting information.

The view that algorithms cannot create information but only shuffle it around is longstanding and well-established. Over 50 years ago, Leon Brillouin, a pioneer in information theory, made that very point: "The [computing] machine does not create any new information, but it performs a very valuable transformation of known information."⁴⁶ Nobel prize-winning biologist Peter Medawar made the same point in the 1980s: "No process of logical reasoning—no mere act of mind or computer-programmable operation—can enlarge the information content of the axioms and premises or observation statements from which it proceeds."⁴⁷

To see that Darwinian processes produce information only in the weaker sense of shuffling preexisting information rather than in the stronger sense of generating it from scratch, we need some way of measuring the information outputted by evolutionary processes. This we have provided in a series of technical papers forthcoming in the computational intelligence literature.⁴⁸ Yet the basic idea is straightforward. Consider again Dawkins's WEASEL. What allowed his evolutionary algorithm to converge so quickly on the target phrase METHINKS•IT•IS•LIKE•A•WEASEL is that a fitness function gauging distance from that phrase was embedded in the algorithm (indeed, the very target phrase was itself stored in the algorithm). But in that case, fitness functions gauging distance from any other string of letters and spaces could just as well have been substituted for the one Dawkins used; and with those other fitness functions, the algorithm could have converged on any sequence whatsoever.

So the target sequence METHINKS•IT•IS•LIKE•A•WEASEL initially had very small probability p (roughly 1 in 10^{40}) of arising by pure chance from a single query; and it has probability q (close to 1) of arising from Dawkins's evolutionary algorithm in a few dozen queries. But that algorithm requires a precisely specified fitness function that gauges distance from a target sequence, and such a fitness function can be built on any sequence of 28 letters and spaces (and not just on METHINKS•IT•IS•LIKE•A•WEASEL). So how many such fitness functions exist? Roughly 10^{40} . And what's the probability of finding Dawkins's fitness function (which gauges distance from METHINKS•IT•IS•LIKE•A•WEASEL) among all these other possible fitness functions? Roughly 1 in 10^{40} .

Indeed, without grammatico-semantic structures on this space of fitness functions (and if such structures exist in the background and constrain the choice of fitness functions, where did that information come from?), the probability distribution on this space is uniform and the probability itself is p , or roughly 1 in 10^{40} . So the gain in probability associated with Dawkins's evolutionary algorithm readily finding the target sequence (i.e., q in place of p) is offset by the improbability of finding the fitness function that gauges distance from that sequence (i.e., p). Dawkins's algorithm, far from explaining how METHINKS•IT•IS•LIKE•A•WEASEL could be produced with high probability, simply raises the new problem of how one overcomes the low

probability of finding the right fitness function for his algorithm. Dawkins has thus filled one hole by digging another.⁴⁹

Simulations such as Dawkins's WEASEL, Adami's AVIDA, Ray's Tierra, and Schneider's *ev* appear to support Darwinian evolution, but only for lack of clear accounting practices that track the information smuggled into them. These programs capitalize on ignorance of how information works. The information hidden in them can be uncovered through a quantity we call *active information*. Active information is to informational accounting what the balance sheet is to financial accounting. Just as the balance sheet keeps track of credits and debits, so active information keeps track of inputs and outputs of information, making sure that they receive their proper due. Information does not magically materialize. It can be created by intelligence or it can be shunted around by natural forces. But natural forces, and Darwinian processes in particular, do not create information. Active information enables us to see why this is the case.

Active information tracks the difference in information between a baseline blind search, which we call the *null search*, and a search that does better at finding the target, which we call the *alternative search*. Consider therefore a search for a target T in a search space Ω (assume for simplicity that Ω is finite). The search for T begins without any special structural knowledge about the search space that could facilitate locating T . Bernoulli's principle of insufficient reason therefore applies and we are in our epistemic rights to assume that the probability distribution on Ω is uniform, with probability of T equal to $p = |T|/|\Omega|$, where $|*|$ is the cardinality of $*$.⁵⁰ We assume that p is so small that a blind or null search over Ω for T (i.e., a search for T by uniform random sampling of Ω) is extremely unlikely to succeed. Success demands that in place of a blind search, an alternative search S be implemented that succeeds with a probability q that is considerably larger than p .

Whereas p gauges the inherent difficulty of locating the target T via a blind search, q gauges the difficulty of locating T via the alternative search S . The question then naturally arises how the blind or null search that locates T with probability p gave way to the alternative search S that locates T with probability q . In the WEASEL, for instance, Dawkins starts with a blind search whose probability of success in one query is roughly 1 in 10^{40} . This is p . He then implements an alternative search (his evolutionary algorithm) whose probability of success in a few dozen queries is close to 1. This is q .

Dawkins leaves the discussion hanging, as though having furnished an evolutionary algorithm that locates the target phrase with high probability (which we are calling S), he has demonstrated the power of Darwinian processes. But in fact all he has done is shifted the problem of locating the target elsewhere, for as we showed earlier in this section, the fitness function he used for his evolutionary algorithm had to be carefully chosen and constituted 1 of 10^{40} (i.e., p) such possible fitness functions. Thus, in furnishing an alternative search whose probability of success is q , he incurred a probability cost p of finding the right fitness function, which coincides (not coincidentally) with the original improbability of the null search finding the target. The information problem that Dawkins purported to solve is therefore left completely unresolved!

In such discussions, it helps to transform probabilities to information measures (note that all logarithms in the sequel are to the base 2). We therefore define the *endogenous information* I_Ω as $-\log(p)$, which measures the inherent difficulty of a blind or null search in exploring the underlying search space Ω to locate the target T . We then define the *exogenous information* I_S as $-\log(q)$, which measures the difficulty of the alternative search S in locating the target T . And finally we define the *active information* I_+ as the difference between the endogenous and exogenous information: $I_+ = I_\Omega - I_S = \log(q/p)$. Active information therefore measures the

information that must be *added* (hence the plus sign in I_+) on top of a null search to raise an alternative search's probability of success by a factor of q/p .

Dawkins's WEASEL, Adami's AVIDA, Ray's Tierra, and Schneider's *ev* are alternative searches. As such, they improve on a null search by increasing the probability of successfully locating targets. In information-theoretic terms, these simulations replace I_Ω with I_S . The endogenous information I_Ω is large, indicating the extreme difficulty of finding the target with the blind or null search. The exogenous information I_S , by contrast, is much smaller (possibly zero), indicating the relative ease of finding the target with an alternative search S . In replacing I_Ω with I_S , these simulations fail to account for the difference in these quantities. In other words, they fail to account for the active information I_+ .

6 Three Conservation of Information Theorems

Active information, though easy to define mathematically, captures a profound truth. We can begin to see this with the following example. Consider an extremely large space Ω that we must search in order to find an extremely tiny target T . In other words, this is a classic needle-in-the-haystack problem. The probability of T with respect to the null search is therefore extremely small, call it p . We might picture Ω as all the dry land on earth and T as a treasure chest buried two feet underground. Next, consider an alternative search S for T conducted in the subspace Ω' of Ω (i.e., $T \subset \Omega' \subset \Omega$) for which the probability of successfully searching for T within Ω' is q (which we assume to be much larger than p). Thus, we might picture Ω' as some small Pacific island (say Bora Bora) on which the treasure T is buried. In this case, $I_\Omega = -\log(p) = -\log \mathbf{P}(T|\Omega)$ and $I_S = -\log(q) = -\log \mathbf{P}(T|\Omega')$. (Note that the conditional probability $\mathbf{P}(T|\Omega)$ is by definition $\mathbf{P}(T \cap \Omega)/\mathbf{P}(\Omega)$, which, since T is contained in Ω and $\mathbf{P}(\Omega) = 1$, is $\mathbf{P}(T) = p$; likewise, because T is contained in Ω' , $\mathbf{P}(T|\Omega') = \mathbf{P}(T \cap \Omega')/\mathbf{P}(\Omega') = \mathbf{P}(T)/\mathbf{P}(\Omega') = q$.)

The search has now become much easier, reduced from all the dry land on earth to Bora Bora. But what enabled the search to become easier? Simply listing the endogenous and exogenous information leaves a crucial question unexplained. Indeed, it is not enough to know that the null search has the very high difficulty level $I_\Omega = -\log(p)$, but that by choosing an appropriate subspace we switch to an alternative search S whose difficulty level $I_S = -\log(q)$ is much lower. The key question that needs to be answered is how we knew to switch the search for T from Ω to the subspace Ω' . In other words, how did we know that of all places on earth where the treasure might be hidden, we needed to look on Bora Bora?

Within the larger space Ω , the subspace Ω' has probability $\mathbf{P}(\Omega'|\Omega) = p/q$ (this follows from $T \subset \Omega' \subset \Omega$, $\mathbf{P}(T|\Omega) = p$, and $\mathbf{P}(T|\Omega') = q$). So the information associated with this subspace is $-\log \mathbf{P}(\Omega'|\Omega) = \log(q/p) = I_+$. Accordingly, when an alternative search improves performance over a null search by reducing the original search space, that improved performance must be paid for by the active information associated with knowing which reduction of the original space (i.e., which subspace) contains the target and therefore should be chosen. Indeed, what prompted us to come up with Ω' in the first place, and how did we know that it contains T ? Active information assigns a cost to any knowledge that enables us to answer this question.

The preceding example is a special case of the following theorem.

Conservation of Information Theorem (function-theoretic version). Let T be a target in Ω . Assume that Ω is finite and nonempty, and that $p = |T|/|\Omega|$ (which we take to be extremely small). The endogenous information is therefore $I_\Omega = -\log(p)$. Next, let Ω' be another nonempty finite space, φ be a function that maps Ω' to Ω , and $T' = \{y \in \Omega' \mid \varphi(y) \in T\}$. Or, in standard set-

theoretic notation, $T' = \varphi^{-1}(T)$. Define $q = |T'|/|\Omega'|$ (which we take to be considerably bigger than p). Given a null search for T' in Ω' , φ induces an alternative search S for T in Ω . The exogenous information is therefore $I_S = -\log(q)$. Next, define \mathcal{F} as the set of all functions from Ω' to Ω and \mathcal{T} as the set of all functions ψ in \mathcal{F} such that $|\psi^{-1}(T)|/|\Omega'| \geq q$ (i.e., each such ψ maps at least as many elements of Ω' to T as φ). Then $|\mathcal{T}|/|\mathcal{F}| \leq p/q$, or equivalently the (higher-order) endogenous information associated with finding \mathcal{T} in \mathcal{F} , i.e., $-\log(|\mathcal{T}|/|\mathcal{F}|)$, is bounded below by the (lower-order) active information $I_+ = \log(q/p)$.

Remarks. To see that this theorem includes the preceding example as a special case, let Ω' be a subset of Ω and let φ be the embedding function that takes each element in the subset to itself in the superset.

This theorem illustrates that the inherent difficulty of a search never goes away. The original (null) search on Ω for T is difficult, characterized by the small probability p of finding the target. We might then imagine that if only we could conduct an easier null search on a space Ω' for a target T' (the probability of success being $q \gg p$) and if only there were some straightforward way to translate target elements of this alternative search to target elements of the original search, the difficulty of the original search would dissolve. But the translation scheme that connects the two searches (in this case, the function φ) resides in its own higher-order search space (the space of functions between the two search spaces), which includes lots of other translation schemes. And how do we search among them?

According to this theorem, finding a translation scheme that maintains the same level of performance as the function φ requires an expenditure of information at least that of the active information. The null search has an inherent difficulty $I_\Omega = -\log(p)$. The alternative search, by utilizing φ , has a significantly reduced difficulty $I_S = -\log(q)$. But finding the φ which enables this reduced difficulty has itself difficulty no less than $I_\Omega - I_S = I_+ = \log(q/p)$. So constructing the alternative search does nothing to make the original problem easier and quite possibly makes matters worse. It's as though one can never fool the original search.

The significance of this theorem, and of conservation of information theorems generally, is that they track the information that was externally applied to augment the probability of successfully searching for T and show that this information is bounded below by the active information. In other words, these theorems show that the improved ease of search, as represented by I_S superseding I_Ω , can only be purchased, in informational terms, at the cost of $I_+ = I_\Omega - I_S$. Conservation of information therefore shows that any improvement in the effectiveness of an alternative search over a null search is not a free lunch. Moreover, payment can never dip below the active information. In fact, *active information represents the optimal price for improving a search*.

Finally, a technical remark needs to be made about how we are representing searches. Right now, given our statement of this theorem, it appears that any search consists of exactly one query. But, in general, a search consists of up to m queries, where m is the maximum number of queries that are practicable (recall section 4). Yet, by redefining the search so that its space consists of the m -fold Cartesian product of the original search space and by redefining the target as the set of all m -tuples from this m -fold product for which at least one coordinate is in the original target, it's possible to reformulate any m -query search so that it is mathematically equivalent to a single-query search on this Cartesian product. In consequence, single-query searches such as appear in this theorem entail no loss of generality.

Proof. Let $\Omega = \{x_1, x_2, \dots, x_K, x_{K+1}, \dots, x_M\}$ so that $T = \{x_1, x_2, \dots, x_K\}$, and let $\Omega' = \{y_1, y_2, \dots,$

$y_L, y_{L+1}, \dots, y_N\}$ so that $T' = \{y_1, y_2, \dots, y_L\}$. Then $p = K/M$ and $q = L/N$ and $|\mathcal{F}| = M^N$. From the binomial theorem it then follows that the number of functions in \mathcal{F} that map L elements of Ω' into T and the remaining elements of Ω' into $\Omega \setminus T$ is

$$\binom{N}{L} K^L (M - K)^{N-L}.$$

From this it follows in turn that the number of functions in \mathcal{F} that map L or more elements of Ω' into T and the remaining elements of Ω' into $\Omega \setminus T$

$$\sum_{i \geq L} \binom{N}{i} K^i (M - K)^{N-i}.$$

Now, if we divide this number by the total number of elements in \mathcal{F} , i.e., by M^N , we get

$$\sum_{i \geq L} \binom{N}{i} p^i (1 - p)^{N-i},$$

which is a cumulative distribution for a binomial random variable with parameters N and p . It is also the probability of \mathcal{T} . Since the mean for such a random variable is Np and since $q = L/N$, it follows that

$$\begin{aligned} \sum_{i \geq L} \binom{N}{i} p^i (1 - p)^{N-i} &= \frac{L}{N} \sum_{i \geq L} \binom{N}{i} p^i (1 - p)^{N-i} \\ &= \frac{1}{N} \sum_{i \geq L} L \binom{N}{i} p^i (1 - p)^{N-i} \\ &\leq \frac{1}{N} \sum_{i \geq 0} i \binom{N}{i} p^i (1 - p)^{N-i} \\ &= \frac{Np}{N} \\ &= \frac{p}{q}. \end{aligned}$$

It follows that $-\log(|\mathcal{T}|/|\mathcal{F}|)$ is bounded below by the active information $I_+ = \log(q/p)$. This proves the theorem.

The conservation of information theorem just proved is perhaps the most basic of all the conservation of information theorems. It shows how, in constructing an alternative search that improves on the average performance over a null search, one must pay for the improvement with an amount not below the active information. All conservation of information theorems take this form. Still, one might wonder whether less expensive ways exist for improving search performance, ways that circumvent conservation of information.

Consider, for instance, Steven Pinker's proposal that the mind is a coordinated assembly of computational modules.⁵¹ We might therefore represent the activity of the mind in performing an information-producing act as a search through a Cartesian product of several (finite) search spaces: $\Omega_1 \times \Omega_2 \times \dots \times \Omega_n$, which we denote by Ω . The target then, let us say, consists of some set $T = T_1 \times T_2 \times \dots \times T_n$ where each T_i is a nonempty subset of Ω_i . We assume that T has extremely

small uniform probability in Ω , which we denote by p . This is the probability of a successful null search. Note that the uniform probability on the Cartesian product is the product of the uniform probabilities, and therefore that the uniform probability of T is the product of the uniform probabilities of the T_i s.

Now Pinker, as a materialist and computational reductionist, would refuse to explain the mind's success in locating T as some fluke that managed, by sheer dumb luck, to find a needle in the haystack. The chance of the gaps is as scientifically unacceptable as the god of the gaps. Thus instead he would want to attribute any such success to the coordination of computational modules where the computational modules and their coordination are the result of a Darwinian evolutionary process. Think of each of these modules, denoted by M_i , as working on its respective space Ω_i to find a target element in T_i . In other words, each M_i delivers an element t_i in T_i and together, as a result of their coordination, these modules induce an alternative search S that delivers the n -tuple (t_1, t_2, \dots, t_n) in the target T . In this wise, computational modules representing basic mental functions can be seen to successfully search Ω for T . Pinker's modular theory of mind would therefore seem well on the road to vindication.

In fact, such an approach raises far more difficulties than it resolves. What, exactly, are these modules and how, specifically, are they coordinated? Pinker never says. But even if he did or could say, the information problem that these modules are supposed to resolve remains as unresolved as ever. To see this, consider that the success of these modules in locating the target depends on them increasing the probability of success well beyond the minuscule probability p for success of the null search, characterized by a uniform probability U on Ω . Pinker's modules, therefore, induce an alternative search S whose probability of success, call it q , is much bigger than p . Thus, his modules replace the uniform probability U with a new probability distribution on Ω , call it μ , that assigns probability q to T .

But where did this μ come from? Did it just magically materialize? No, it resides in the space of probability measures on Ω , and in that space it has to be found and identified. But how probable is it that we can find, in this space of probability measures, a probability measure (more often called a "probability distribution") at least as effective as μ at locating T ? As the next theorem demonstrates, the (higher-order) probability of finding a probability distribution ν that's at least as effective as μ at locating T is less than or equal to p/q .

Conservation of Information Theorem (measure-theoretic version). Let T be a target in Ω . Assume Ω is finite and T is nonempty. Let U denote the uniform probability distribution on Ω and let $p = |T|/|\Omega| = U(T)$ (which we take to be extremely small). The endogenous information is therefore $I_\Omega = -\log(p)$. Next, let μ be a probability distribution on Ω such that $q = \mu(T)$ (which we take to be considerably bigger than p). Suppose that μ characterizes the probabilistic behavior of an alternative search S . The exogenous information is therefore $I_S = -\log(q)$. Next, let \mathcal{M} denote the set of all probability distributions on Ω and \mathcal{T} be the set of probability distributions ν in \mathcal{M} such that $\nu(T) \geq q$ (i.e., each such ν assigns at least as much probability to T as μ —each such ν therefore represents a search that's at least as effective at locating T as μ). Then the (higher-order) uniform probability of \mathcal{T} in \mathcal{M} , which may be denoted by $U^*(\mathcal{T})$, is less than or equal to p/q . Equivalently, the (higher-order) endogenous information associated with finding \mathcal{T} in \mathcal{M} , i.e., $-\log(U^*(\mathcal{T}))$, is bounded below by the (lower-order) active information $I_+ = -\log(U(T)) + \log(\mu(T)) = \log(q/p)$.

Proof. Let $\Omega = \{x_1, x_2, \dots, x_K, x_{K+1}, \dots, x_N\}$ so that $T = \{x_1, x_2, \dots, x_K\}$. Then $p = K/N$. Next, given that μ is a probability distribution on Ω , it follows that μ has the form

$$\sum_{i=1}^N a_i \delta_{x_i},$$

where each a_i is nonnegative, the a_i s together sum to 1, and each δ is a point mass (assigning probability 1 to the corresponding x_i). Indeed, each element of \mathcal{M} has this form. It follows that \mathcal{M} has the geometric structure of an $(N-1)$ -dimensional simplex consisting of all convex combinations of N nonnegative real numbers. Moreover, its uniform probability is given by a normalized Lebesgue measure.

Since $\mu(T) = q$, it follows that

$$\sum_{i=1}^K a_i = q.$$

Moreover, any distribution ν in \mathcal{T} of the form

$$\sum_{i=1}^N b_i \delta_{x_i}$$

satisfies

$$\sum_{i=1}^K b_i \geq q.$$

From these facts it now follows that the uniform probability \mathbf{U}^* of \mathcal{T} in \mathcal{M} is given by the following expression:⁵²

$$\frac{\Gamma(N)}{\Gamma(N(1-p))\Gamma(Np)} \int_0^{1-q} t^{N(1-p)-1} (1-t)^{Np-1} dt.$$

This last expression describes a cumulative beta distribution with first parameter $r = N(1-p)$ and second parameter $s = Np$. Integration by substitution shows that this expression can be rewritten as

$$\frac{\Gamma(N)}{\Gamma(Np)\Gamma(N(1-p))} \int_q^1 t^{Np-1} (1-t)^{N(1-p)-1} dt,$$

which describes a cumulative beta distribution with first parameter $r = Np$ and second parameter $s = N(1-p)$. It is well known that the mean for this distribution is $r/(r+s)$.⁵³ In consequence,

$$\begin{aligned} \frac{\Gamma(N)}{\Gamma(Np)\Gamma(N(1-p))} \int_q^1 t^{Np-1} (1-t)^{N(1-p)-1} dt &= \frac{\Gamma(N)}{\Gamma(Np)\Gamma(N(1-p))} \int_q^1 \frac{q}{q} \cdot t^{Np-1} (1-t)^{N(1-p)-1} dt \\ &= \frac{1}{q} \cdot \frac{\Gamma(N)}{\Gamma(Np)\Gamma(N(1-p))} \int_q^1 q \cdot t^{Np-1} (1-t)^{N(1-p)-1} dt \\ &\leq \frac{1}{q} \cdot \frac{\Gamma(N)}{\Gamma(Np)\Gamma(N(1-p))} \int_0^1 t \cdot t^{Np-1} (1-t)^{N(1-p)-1} dt \\ &= \frac{1}{q} \cdot \frac{Np}{Np + N(1-p)} \\ &= \frac{p}{q}. \end{aligned}$$

It follows that $-\log(\mathbf{U}^*(\mathcal{T}))$ is bounded below by the active information $I_+ = \log(q/p)$. This proves the theorem.

Conservation of information is also implicit in a family of mathematical results known as NFL theorems (i.e., no free lunch theorems). Think of an evolutionary search as treating fitness functions as a variable, with each fitness function providing a different representation of selective advantage for the environment. A typical NFL theorem then states that the average performance of an evolutionary search across fitness functions does not exceed blind search. Conservation of information's connection to NFL now becomes evident as soon as one inquires into what enables certain fitness functions to induce alternative (evolutionary) searches that are so much better than blind search, especially given that (by NFL) averaging across fitness functions does no better than blind search. According to conservation of information, active information is required to locate the fitness functions that render evolutionary search effective. The following theorem underscores the connection between conservation of information and no free lunch.

Conservation of Information Theorem (fitness-theoretic version). Let T be a target in Ω . Assume Ω is finite and T is nonempty. Let \mathbf{U} denote the uniform probability distribution on Ω and let $p = |T|/|\Omega| = \mathbf{U}(T)$ (which we take to be extremely small). The endogenous information is therefore $I_\Omega = -\log(p)$. Next, let \mathcal{F} denote a finite collection of fitness functions on Ω and let \mathbf{S}_Ω denote the symmetric group on Ω (i.e., all permutations of this set). Without loss of generality, assume that any f in \mathcal{F} only takes values in $\{0, 1/M, 2/M, \dots, (M-1)/M, 1\}$ for some large fixed M and that \mathcal{F} includes all such f . \mathcal{F} is therefore closed under the symmetric group \mathbf{S}_Ω , i.e., for any f in \mathcal{F} and any σ in \mathbf{S}_Ω , $f \circ \sigma$ is also in \mathcal{F} . Suppose further that any f in \mathcal{F} induces a probability distribution \mathbf{P}_f on Ω (corresponding to an alternative search). Assume that each such \mathbf{P}_f satisfies the following invariance property: for any σ in \mathbf{S}_Ω and $A \subset \Omega$, $\mathbf{P}_{f \circ \sigma}(\sigma^{-1}A) = \mathbf{P}_f(A)$. An NFL result then follows:

$$\frac{1}{|\mathcal{F}|} \sum_{f \in \mathcal{F}} \mathbf{P}_f(T) = \mathbf{U}(T).$$

Assume next that g in \mathcal{F} is such that $q = \mathbf{P}_g(T)$ (which we take to be considerably bigger than p) and that g induces an alternative search S for which the exogenous information is $I_S = -\log(q)$. Let \mathcal{T} denote the set of all fitness functions h in \mathcal{F} such that $\mathbf{P}_h(T) \geq q$ (i.e., each such h induces a probability distribution that assigns at least as much probability to T as \mathbf{P}_g ; equivalently, each such h induces an alternative [or evolutionary] search at least as effective at locating T as S). Then, the (higher-order) uniform probability of \mathcal{T} in \mathcal{F} , i.e., $|\mathcal{T}|/|\mathcal{F}|$, which may be denoted by $\mathbf{U}^*(\mathcal{T})$, is less than or equal to p/q . Equivalently, the (higher-order) endogenous information associated with finding \mathcal{T} in \mathcal{F} , i.e., $-\log(\mathbf{U}^*(\mathcal{T}))$, is bounded below by the (lower-order) active information $I_+ = -\log(\mathbf{U}(T)) + \log(\mathbf{P}_g(T)) = \log(q/p)$.

Remarks. This theorem attempts to characterize the informational properties of evolutionary algorithms that use fitness functions f in \mathcal{F} to search for possible targets in Ω (the target of interest being T). The assumption that the probability distributions \mathbf{P}_f induced by these fitness functions are invariant under a group action (in this case, the symmetric group \mathbf{S}_Ω) is customary with such theorems: it ensures that fitness is not correlated with prior knowledge of the target. As Joseph Culberson puts it: “Evolutionary algorithms (EAs) are often touted as ‘no prior knowledge’ algorithms. This means that we expect EAs to perform without special information

from the environment. Similar claims are often made for other adaptive algorithms.”⁵⁴ If f and $f \circ \sigma$ were not equally effective at locating the targets T and $\sigma^{-1}T$ respectively, presumably special information from the environment (and thus prior knowledge not available to the fitness function f) would account for the difference.

This theorem includes a no free lunch result as well as a conservation of information result (whose proof depends on the no free lunch result). The formula

$$\frac{1}{|\mathcal{F}|} \sum_{f \in \mathcal{F}} \mathbf{P}_f(T) = \mathbf{U}(T)$$

expresses a classic instance of NFL: it shows that an evolutionary search, when averaged across fitness functions, does no better at locating a target T than blind or null search (in fact, it shows the two are identical). Note that because of the invariance property of the probability distributions \mathbf{P}_f under the symmetric group \mathbf{S}_Ω , any constant function c over Ω that's in \mathcal{F} induces a probability distribution \mathbf{P}_c that is identical to the uniform probability \mathbf{U} . This follows because $c \circ \sigma = c$ for all σ in \mathbf{S}_Ω .

Proof. Let $\Omega = \{x_1, x_2, \dots, x_K, x_{K+1}, \dots, x_N\}$ so that $T = \{x_1, x_2, \dots, x_K\}$. Then $p = K/N$. Without loss of generality, assume that N is divisible by K so that $K \times L = N$ for some L . If not, simply expand Ω with enough finite elements so that T has probability p and K does divide N . Next, find a set of elements $\sigma_1, \sigma_2, \dots, \sigma_L$ in the symmetric group \mathbf{S}_Ω such that $T_i = \sigma_i^{-1}(T)$ and the T_i s partition Ω with $T_1 = T$. In other words, the T_i s are disjoint, include T , and their union equals all of Ω .

Because the probability distributions induced by the fitness functions f in \mathcal{F} are invariant under the symmetric group \mathbf{S}_Ω , it follows that for any σ in \mathbf{S}_Ω ,

$$\sum_{f \in \mathcal{F}} \mathbf{P}_{f \circ \sigma}(\sigma^{-1}T) = \sum_{f \in \mathcal{F}} \mathbf{P}_f(T).$$

But since as f runs through all the fitness functions in \mathcal{F} , $f \circ \sigma$ likewise runs through all the fitness functions in \mathcal{F} (that's because \mathcal{F} is closed under composition with elements from \mathbf{S}_Ω), it follows that

$$\sum_{f \in \mathcal{F}} \mathbf{P}_{f \circ \sigma}(\sigma^{-1}T) = \sum_{f \in \mathcal{F}} \mathbf{P}_f(\sigma^{-1}T).$$

In consequence, summing all the fitness-induced probability measures over \mathcal{F} is also invariant under \mathbf{S}_Ω . And this in turn implies that for each i ($1 \leq i \leq L$)

$$\frac{1}{|\mathcal{F}|} \sum_{f \in \mathcal{F}} \mathbf{P}_f(\sigma_i^{-1}T) = \frac{1}{|\mathcal{F}|} \sum_{f \in \mathcal{F}} \mathbf{P}_f(T_i) = \frac{1}{|\mathcal{F}|} \sum_{f \in \mathcal{F}} \mathbf{P}_f(T)$$

is constant. But since the T_i s partition Ω with $T_1 = T$, it follows that

$$\sum_{i=1}^L \frac{1}{|\mathcal{F}|} \sum_{f \in \mathcal{F}} \mathbf{P}_f(T_i) = 1$$

and therefore that

$$\frac{1}{|\mathcal{F}|} \sum_{f \in \mathcal{F}} \mathbf{P}_f(T) = \frac{1}{L} = \frac{K}{K \cdot L} = \frac{K}{N} = p = \mathbf{U}(T).$$

This establishes the first main claim of the theorem.

As for $\mathbf{U}^*(\mathcal{T}) = |\mathcal{T}|/|\mathcal{F}|$ being bounded above by p/q , suppose, for the sake of a contradiction, that $|\mathcal{T}| > (p/q)|\mathcal{F}|$. That would mean that in the sum

$$\sum_{f \in \mathcal{F}} \mathbf{P}_f(T),$$

more than p/q of the summation elements are greater than or equal to $\mathbf{P}_g(T) = q$. But that in turn would mean that the normalized sum

$$\frac{1}{|\mathcal{F}|} \sum_{f \in \mathcal{F}} \mathbf{P}_f(T).$$

would have to be strictly greater than p . But we just established that it is equal to p .

From this contradiction it follows that $\mathbf{U}^*(\mathcal{T}) = |\mathcal{T}|/|\mathcal{F}|$ is indeed less than or equal to p/q and therefore that $-\log(|\mathcal{T}|/|\mathcal{F}|)$ is bounded below by the active information $I_+ = \log(q/p)$. This proves the theorem.

This fitness-theoretic conservation of information theorem is more significant than it might at first appear. One might think that its applicability is limited because it was formulated in such general and seemingly unrealistic terms. What search space, for instance, allows for all possible permutations? Most don't. Yet, insofar as they don't, it's because they exhibit structures that constrain the permissible permutations. Such constraints, however, bespeak the addition of active information. Consider, for instance, that most evolutionary algorithms are used to search not a completely unstructured space Ω but an m -fold Cartesian product space Ω'^m , each factor Ω' signifying a single query in an m -query search.

In that case, permissible permutations must not scramble query-order or vary from one factor to the next but rather must act the same way on Ω' across factors (if a permutation moves a point from one place to another in one factor, then it must do the same in the other factors). Thus, in place of the full symmetric group on Ω'^m , whose cardinality is the factorial of $|\Omega'|^m$ (i.e., $|\Omega'|^m!$), the relevant group action would be the symmetric group on Ω' , whose cardinality is the factorial of $|\Omega'|$ (i.e., $|\Omega'|!$). But in resulting ratio, $|\Omega'|!/|\Omega'|^m!$, the denominator completely overwhelms the numerator. It therefore represents a huge input of active information; indeed, unless these spaces are very very small, this ratio will be much less than p/q .

In a similar vein, one might want to constrain the fitness functions. Thus, one might think that fitness ought to vary gradually with some underlying metric structure on the search space. But where does such a metric structure come from? And how much does it reduce the full space of fitness functions \mathcal{F} ? If it reduced \mathcal{F} to some smaller space of fitness functions \mathcal{F}' , then $|\mathcal{F}'|/|\mathcal{F}|$ represents a further addition of active information, as does the metric structure on the underlying search space Ω (how many other metric or topological structures were possible and what led to this one taking effect rather than the others?).

Christian Igel and Marc Toussaint have, for instance, argued that NFL theorems are unrealistic because they focus on fitness functions closed under permutation. They suggest that in realistic problems the focus should instead be on classes of fitness functions that are not closed under permutation.⁵⁵ All such focusing and constraining, however, imparts active information.

Moreover, once such constraints on permissible fitness functions or permissible permutations are fleshed out to the point where we can calculate how much active information was imparted, we find invariably that conservation of information is preserved.

Indeed, all the conservation theorems listed in this section (and those we know of otherwise) give active information as the extreme lower bound on the amount of information that must be imparted to an alternative search to improve it by a given amount over a null search. Take the measure-theoretic conservation of information theorem proved earlier in this section. We have proved a more precise result showing that in the search for probability distributions that represent successful (lower-order) searches, the information cost need not simply be bounded below by the active information $I_+ = \log(q/p)$ (as we showed in the theorem proved in this section) but can grow exponentially in I_+ , indicating runaway informational inflation as one attempts to account for the information required to render an alternative search successful.⁵⁶

7 The Law of Conservation of Information

Laws of science are supposed to be universal in scope, hold with unfailing regularity, and find support from a wide array of facts and observations. We submit that conservation of information is such a law. It may be formulated as follows:

The Law of Conservation of Information (LCI). Any search that proportionately raises the probability of locating a target by q/p with respect to blind search requires in its formation an amount of information not less than the active information $I_+ = \log(q/p)$.

In short, raise the probability of successful search by a factor of q/p , incur an information cost of $\log(q/p)$. The rest of this section consists of bullet points devoted to elucidating this law.

General Setup. The general setup for LCI is as follows. A null search B , which is blind, sets a probabilistic baseline p for searching the space Ω for a target T . Think of B as an Ω -valued random variable that induces a uniform probability distribution on Ω . Regardless of what we know or don't know about T , we can always implement B and therefore do at least as good as B in searching Ω for T . The question is how much better can we do than B . In practice, the probability p of B locating T is so small that B stands no reasonable chance of success. We therefore look to an alternative search S whose probability q of locating T is significantly larger than p . But where did S come from? S did not magically materialize. Rather, it had to be brought into existence through some formative process. LCI states that the formation of S required an investment of information not less than $I_+ = \log(q/p)$.

A Family of Theorems. LCI receives support from conservation theorems of the sort proved in the last section. Such theorems confirm that the information needed to improve a search so that its probability of successfully locating a target increases by a factor of q/p is at least $I_+ = \log(q/p)$. Even so, LCI is not itself a mathematical theorem. It says that in any circumstance where a null search gives way to an improved alternative search, at least I_+ needs to be imparted to the alternative search. But the specific forms by which null and alternative searches can be instantiated is so endlessly varied that no single mathematical theorem can cover all contingencies.

That's why, in the last section, we offered three substantially different types of conservation of information theorems. Instead of LCI falling under one general mathematical theorem, LCI

says that for any mathematically specific situation in which a blind search gives way to an improved alternative search, a mathematical theorem exists showing that the alternative search required at least I_+ to be formed. Thus, instead of LCI constituting a theorem, it characterizes situations in which we may legitimately expect to prove a conservation of information theorem. LCI might therefore be viewed as a family of theorems sharing certain common features.

No Free Lunch Theorems. In the 1990s, William Macready and David Wolpert proved several theorems to which they gave the descriptive tag “No Free Lunch” or “NFL.”⁵⁷ These theorems showed how evolutionary searches, when averaged across fitness functions, did no better at locating targets than blind (or null) search. Their work had precedent. Cullen Schaffer’s Law of Conservation of Generalization Performance, proved slightly earlier, compared a learner who can consistently achieve “at least mildly better-than-chance performance” to “a perpetual motion machine.”⁵⁸ Schaffer’s point was that just as perpetual motion machines constitute a thermodynamic impossibility, a learner who consistently outperforms chance (even if only mildly so) constitutes an algorithmic impossibility. He elaborated,

An essential property of the problem of inductive generalization is that it admits no general solution. An algorithm that is good for learning certain sets of concepts must necessarily be bad for learning others. Moreover, no algorithm strictly dominates any other. If two learners differ in generalization performance, there must be problems for which each is superior to the other. As a consequence, *every* algorithm is appropriate in some sphere of application and each is thus, in a sense, worthy of study.⁵⁹

After Macready and Wolpert’s initial work in this area, other researchers proved additional NFL theorems. Yu-Chi Ho and David Pepyne offered a simplified approach to these theorems.⁶⁰ Thomas English formulated an information-theoretic NFL theorem, which he described explicitly in terms of “conservation of information,” calling it a “law” as well as a “principle” and relating it to the work of Schaffer. English’s approach underscored the futility of attempting to design a generally superior optimizer without problem-specific information about a search.

Work on NFL theorems, in focusing on average performance of search algorithms, leads to a natural follow-up question, namely, what are the informational costs associated with finding algorithms that in practice are better than others? NFL is a great leveler of search algorithms. It is counterintuitive because we know that some search algorithms are much better than others *at specific tasks*. The Law of Conservation of Information, as we develop it here, presupposes the NFL theorems and then quantifies the informational costs that makes some algorithms better than others at performing specific tasks. According to LCI, improved performance of an algorithm at performing a specific task is purchased at the cost of active information. Simply put, NFL says there is a cost for effective search; LCI calculates the cost.

Church-Turing Analogue. The Law of Conservation of Information, in both scope and function, parallels the Church-Turing Thesis of theoretical computer science.⁶¹ The Church-Turing Thesis is a deep claim in mathematical logic about the nature of computation. LCI functions within the theory of intelligent design as the Church-Turing Thesis functions within the theory of computation. According to this thesis, given a procedure that is informally computable (i.e., that is finitely specifiable and rule-based), it can be coded as an algorithm running on a Turing machine. According to LCI, given an alternative search that does better than blind search, the alternative search was purchased at an information cost no less than I_+ . The task of the design theorist in that case is to “follow the information trail” and show where the information that this

search outputs in locating a target was first inputted (much as the task of the computer scientist is to show how some procedure that is informally computable can be explicitly formulated as an algorithm capable of being run on a Turing machine).

It follows that there is no—and indeed can be no—strict mathematical proof of either the Church-Turing Thesis or LCI. Nonetheless, the two are subject to independent verification. With the Church-Turing Thesis, the challenge is to show that any operation that is informally computable is also, upon a close mathematical analysis, formally computable (to date, this has always been confirmed). Likewise with LCI, the challenge is to find where the information that improves an alternative search over a blind search was inputted and then to calculate the amount of this information, showing that it does not fall below $I_+ = \log(q/p)$ (to date, this has also always been confirmed).

Probability Distributions vs. Probabilistic Outcomes. Much of the power of LCI comes from its focus not on particular probabilistic outcomes but on probability distributions. LCI is not saying that certain unlikely outcomes or events do not, or cannot, happen. Rather, it is saying that certain types of probability distributions cannot obtain without the addition of information. LCI begins with the observation that different probability distributions are associated with different searches (distributions which in many instances can be given precise mathematical characterizations based on theoretical or empirical considerations). These distributions can assign very different numbers to the probability of successfully concluding a search. According to LCI, any such differences between an alternative and a null search must be accounted for in terms of the information put into forming the alternative search.

Not Computational (or Kolmogorov) Complexity. Given an alternative search that improves on a null search, we might think that simply by inspecting a representation of the alternative search, it would be possible to calculate the information cost that was incurred in its formation. And we would be wrong. When a search is represented computationally, the complexity of its representation may be irrelevant to the information required for its formation. With computational representations of searches, their complexity often depends on idiosyncrasies of the underlying computational environment. Imagine searches whose targets are sequences 1,000-bits in length. These searches might use fitness functions that gauge distance from the target by counting the number of bits that fail to match up (i.e., they might use the Hamming distance).

Now imagine two possible searches in this space of 1,000-bit sequences. In the first, the fitness function gauges distance from a 1,000-bit sequence of high Kolmogorov complexity (i.e., the sequence is highly incompressible computationally, as we would expect with a random sequence of coin tosses). In the other, the fitness function gauges distance from a 1,000-bit sequence of low Kolmogorov complexity (i.e., the sequence is highly compressible, for instance, a sequence of 1,000 zeros). The first fitness function, and thus the search associated with it, will require considerable complexity for its representation. The second one, by contrast, will be much simpler to represent (the main evaluation requiring a line of code that says “add number of bits differing from 0”). And yet both searches, when identified with fitness functions, will require the same amount of information to be located and extracted from the underlying space of fitness functions (compare the fitness-theoretic conservation of information theorem of the last section).

The Search for a Search. In characterizing the information cost that alternative searches incur, LCI treats searches as residing in higher-order search spaces (i.e., spaces each of whose

elements is a search). Notably, the information imparted to improve a search picks out one class of searches (those that with probability q or better locate T) to the exclusion of others (those that with probability less than q locate T). LCI says that searching the original search space Ω for a target T with probability q of success is never more difficult, and possibly much easier, than searching a higher-order search space for a search that, when applied to the lower-order search space Ω , finds T with that same probability.

To see what's at stake in such a "search for a search," imagine that you are on an island with buried treasure. The island is so large that a blind (null) search is highly unlikely to succeed in finding the treasure. Fortunately, you have a treasure map that will guide you to it. But where did you find the treasure map? Treasure maps reside in a library of possible treasure maps. The vast majority of these will not lead to the treasure. How, then, did you happen to find the right map among all these possible treasure maps? What special information did you need to find it? Conservation of information says that the information required to pick out the right map is never less than the information required to locate the treasure directly.

From the vantage of LCI, *searches are as real as the objects being searched*. Just as the existence and formation of those objects must be explained, so too the existence and formation of the searches that locate those objects must be explained. Thus we might say that searches, by residing in a space of searches, are themselves objects to be searched. This implies a hierarchy of searches: the original search, the search for that search, the search for the search for that search, etc. LCI says that as we regress up this search hierarchy, the search problem never becomes easier and may in fact become more difficult. We clarify this point next.

The LCI Regress. Consider again our general setup: A null search B , which is blind, sets a probabilistic baseline p for searching the space Ω for a target T . Because p is so small that B stands no practical possibility of locating T , success in searching for T requires an alternative search S that has probability q of locating T , where q is much larger than p . But that raises the question, how did we find S ? S resides in a higher-order search space, call it $\Omega^{(2)}$, and specifically in a higher-order target $T^{(2)}$ consisting of all searches that have probability at least q of locating the original target T (for consistency let $\Omega^{(1)} = \Omega$ and let $T^{(1)} = T$). But how easy is it to find $T^{(2)}$ in $\Omega^{(2)}$? LCI tells us that doing so requires at least $I_+ = \log(q/p)$. Moreover, once we've found the alternative search S in $T^{(2)}$, we still have to use it to search for the original target T in Ω . This lower-order search has probability q of success, which corresponds to the exogenous information $I_S = -\log(q)$.

Thus, the information required to successfully locate the original target T by first searching a higher-order search space for S and then using S to search for T requires at least $I_+ + I_S = \log(q/p) - \log(q) = -\log(p) = I_\Omega$. This shows that the endogenous information (i.e., inherent difficulty) of using a higher-order search to locate the original target T is at least as great as the original endogenous information. We represent this fact by the inequality

$$I_{\Omega^{(1)}} \leq I_{\Omega^{(2)}} ,$$

where the first term denotes the original lower-order endogenous information of finding T directly (i.e., I_Ω since $\Omega = \Omega^{(1)}$) and the second denotes the higher-order endogenous information of finding T indirectly by first searching $\Omega^{(2)}$ for a higher-order target $T^{(2)}$.

Given LCI, this inequality implies (by mathematical induction) that as we move up the search hierarchy to search for higher-order targets $T^{(3)}$, $T^{(4)}$, etc. within higher-order search spaces $\Omega^{(3)}$, $\Omega^{(4)}$, etc., higher-order endogenous information associated with locating the original target T will never diminish and may well increase. We call this the LCI Regress and write

$$I_{\Omega^{(1)}} \leq I_{\Omega^{(2)}} \leq I_{\Omega^{(3)}} \leq \dots$$

Here $I_{\Omega^{(3)}}$ sums the information needed to locate $T^{(3)}$ plus the information needed to locate $T^{(2)}$ using a search in $T^{(3)}$ plus the information needed to locate the original T using a search in $T^{(2)}$.

So, we may ask, if an alternative search achieves a high probability of locating a target T that is extremely improbable with respect to a null (or blind) search, where did the information that enables the alternative search to be successful come from? From a higher-order search? But, as the LCI Regress shows, a higher-order search requires at least as much information to locate T as any lower-order search. Borrowing from Peter to pay Paul at best maintains, and may even intensify, the debt now owed to Peter. Likewise borrowing from Andrew to pay Peter maintains or intensifies the debt still further. Indeed, borrowing from one lender to pay another does nothing to redress a debt. Where, then, does the LCI Regress end? In fact, it may not end, implying that the information that enables an alternative search to succeed in locating T was present always. Alternatively, it may end because an external information source added the information needed to locate T . One option suggests front-loading of information, the other direct input. Both options evoke intelligent design.

Entropy. The LCI regress suggests a deep connection between the Law of Conservation of Information and the Second Law of Thermodynamics. In *Logic and Information*, Keith Devlin considers the thermodynamic significance of information:

Perhaps *information* should be regarded as (or maybe is) a basic property of the universe, alongside matter and energy (and being ultimately interconvertible with them). In such a theory (or suggestion for a theory, to be more precise), information would be an intrinsic measure of the structure and order in parts or all of the universe, being closely related to entropy (and in some sense its inverse).⁶²

According to the LCI Regress, the information needed for effective search tends to increase as time moves backward (that's because searches proceed in time and the search for a given search necessarily precedes the given search). Yet, according to the Second Law of Thermodynamics, the unavailability of energy for conducting useful work in an isolated non-equilibrium system, as measured by entropy, tends to increase as times moves forward. It seems, then, that information as characterized by the Law of Conservation of Information may be regarded as inverse to entropy: increased information indicates an increased capacity for conducting successful search whereas increased entropy indicates a decreased capacity for doing the work necessary to conduct a search.

Given the paradox of Maxwell's Demon,⁶³ which indicates the ability of an information source to decrease entropy over time, the Law of Conservation of Information might actually prove more basic than the Second Law of Thermodynamics. The title of Leo Szilard's celebrated 1929 paper on Maxwell's Demon is worth recalling here: "On the Decrease of Entropy in a Thermodynamic System by the Intervention of Intelligent Beings."⁶⁴ The information source that for Szilard reverses entropy is intelligent. Likewise, the LCI Regress, as noted in the last bullet point, suggests that intelligence is ultimately the source of the information that accounts for successful search.

8 Applying LCI to Biology

Biology's reception of Darwinism might have been less favorable had scientists heeded Darwin's contemporary John Stuart Mill. In 1843, sixteen years before the publication of Darwin's *Origin of Species*, Mill published the first edition of his *System of Logic* (which by the 1880s had gone through eight editions).⁶⁵ In that work Mill lays out various methods of induction. The one that interests us here is his *method of difference*. Mill described this method as follows:

If an instance in which the phenomenon under investigation occurs, and an instance in which it does not occur, have every circumstance in common save one, that one occurring only in the former; the circumstance in which alone the two instances differ is the effect, or the cause, or an indispensable part of the cause, of the phenomenon.⁶⁶

Essentially this method says that to discover which of a set of circumstances is responsible for an observed difference in outcomes requires finding a difference in the circumstances associated with each outcome. An immediate corollary of this method is that common circumstances cannot explain a difference in outcomes.

Suppose you and a friend have been watching TV, eating popcorn, and lounging on a vibrating couch. Yet your friend is now staggering about bleary-eyed whereas you are moving and seeing just fine. Precisely because the TV, popcorn, and vibrating couch are common to both your experience, they do not explain why your friend is having difficulties and you are doing fine. To explain the difference, you need to find not what's common to your circumstances but what's different. When you remember that your friend also consumed an inordinate amount of alcohol whereas you simply drank water, you've applied Mill's method to explain the difference.

This method, so widely used in everyday life, is crucially relevant to biological evolution. Indeed, it helps bring some sense of proportion and reality to the inflated claims so frequently made on behalf of Darwinian processes. We've already cited (in section 5) Kenneth Miller's overselling of Darwinism, where he claims that "what's needed to drive" increases in biological information is "just three things: selection, replication, and mutation,"⁶⁷ Mill's method of difference gives the lie to Miller's claim. It's easy to write computer simulations that feature selection, replication, and mutation—and *that go absolutely nowhere*. It's also possible to write such simulations that solve interesting problems and produce salient patterns. But because selection, replication, and mutation are common to both such simulations, they can't, as Mill's method makes clear, account for the difference.

Nils Barricelli, writing about computer simulations of evolution as far back as 1961, understood this point:

The *selection* principle of Darwin's theory is not sufficient to explain the evolution of living organisms if one starts with entities having only the property to *reproduce* and *mutate*. At least one more theoretical principle is needed, a principle which would explain how self-reproducing entities could give rise to organisms with the variability and evolutionary possibilities which characterize living organisms.⁶⁸

Barricelli's point here is exactly correct: Miller's holy trinity of selection, replication, and mutation are "not sufficient"; "at least one more theoretical principle is needed." Unfortunately, Barricelli's subsequent proposal for the missing theoretical principle is mistaken. What was his proposal? Symbiogenesis, the idea that organisms or simple replicators can combine to form new, more complex organisms or replicators.⁶⁹

These days, symbiogenesis is most closely associated with the research of biologist Lynn Margulis, who has been working on this topic for several decades.⁷⁰ Although she describes many interesting cases of symbiogenesis, she hasn't shown how this process constitutes a general solution to biology's information problem. Symbiogenesis, whether operating in real or in virtual biology, can at best mix existing traits; it cannot create fundamentally new ones. For instance, when two organisms merge in symbiogenesis, the genetic complement of the newly formed organism is simply the sum of the genes from the original two organisms—no new genes are created. Genuine novelty is therefore beyond symbiogenesis' reach. And yet, genuine novelty is precisely what the steady increase of biological information over the course of natural history exhibits.

Neither Barricelli's nor Margulis's proposals have won the day. The reason they haven't is that symbiogenesis, like genetic mutation, is simply another undirected way of producing raw variation in biological structures. To resolve biology's information problem requires harnessing that variation. For most evolutionary biologists, that's the job of natural selection. Thus Margulis, though an outspoken critic of neo-Darwinism (which locates the source of evolutionary variation in genetic mutation), will admit, "I am definitely a Darwinist." Yet with no apparent sense of irony, she immediately added "I think we are *missing important information* about the origins of variation."⁷¹ The missing ingredient in her account of evolution is in fact staring her in the face: information! Yet rather than focus on the role of information in guiding evolution, Margulis continues to focus on undirected sources of variation. Supplementing Miller's "mutation" with other nonteleological sources of variation such as symbiogenesis does nothing to meet the challenge raised by Mill's method of difference.

The failure of selection, replication, and mutation (or undirected variation more generally—include here symbiogenesis, genetic drift, lateral gene transfer, etc.) to drive evolution is evident not just in computer simulations but also in actual biological experiments. Consider, for instance, Sol Spiegelman's work on the evolution of polynucleotides in a replicase environment. To evolve his polynucleotides, Spiegelman inserted information: the replicase protein was supplied by the investigator from a viral genome as were the activated mononucleotides needed to feed polynucleotide synthesis. Yet even without such investigator interference, which has no analogue in a Darwinian conception of natural history, a deeper problem remains. According to Miller, selection, replication, and mutation (or variation) increase information. Yet Spiegelman demonstrated that even with all these factors at play, information steadily *decreased* over the course of his experiment. Brian Goodwin, in his summary of Spiegelman's work, underscores this point:

In a classic experiment, Spiegelman in 1967 showed what happens to a molecular replicating system in a test tube, without any cellular organization around it. The replicating molecules (the nucleic acid templates) require an energy source, building blocks (i.e., nucleotide bases), and an enzyme to help the polymerization process that is involved in self-copying of the templates. Then away it goes, making more copies of the specific nucleotide sequences that define the initial templates. But the interesting result was that these initial templates did not stay the same; they were not accurately copied. They got shorter and shorter until they reached the minimal size compatible with the sequence retaining self-copying properties. And as they got shorter, the copying process went faster. So what happened with natural selection in a test tube: the shorter templates that copied themselves faster became more numerous, while the larger ones were gradually eliminated. This looks like Darwinian evolution in a test tube. But the interesting result was that this evolution went one way: toward greater simplicity.⁷²

The problem that Spiegelman identified here is merely the tip of the iceberg. Yes, evolution had better be complexity-increasing if it is to deserve all the attention it receives. But complexity also needs to be going somewhere. In fact, in the history of life, increasing complexity has been in the service of building magnificent structures of incredible sophistication and elegance. How could evolution accomplish all this, especially given NFL and LCI? Complexity theorist Stuart Kauffman understands the challenge:

The no-free-lunch theorem says that, averaged over all possible fitness landscapes, no search procedure outperforms any other.... In the absence of any knowledge, or constraint, [read “information”] on the fitness landscape, on average, any search procedure is as good as any other. But life uses mutation, recombination, and selection. These search procedures seem to be working quite well. Your typical bat or butterfly has managed to get itself evolved and seems a rather impressive entity. The no-free-lunch theorem brings into high relief the puzzle. If mutation, recombination, and selection only work well on certain kinds of fitness landscapes, yet most organisms are sexual, and hence use recombination, and all organisms use mutation as a search mechanism, where did these well-wrought fitness landscapes come from, such that evolution manages to produce the fancy stuff around us?⁷³

According to Kauffman, “No one knows.”⁷⁴

Let’s be clear where our argument is headed. We are not here challenging common descent, the claim that all organisms trace their lineage to a universal common ancestor. Nor are we challenging evolutionary gradualism, that organisms have evolved gradually over time. Nor are we even challenging that natural selection may be the principal mechanism by which organisms have evolved. Rather, we are challenging the claim that evolution can create information from scratch where previously it did not exist. The conclusion we are after is that natural selection, even if it is the mechanism by which organisms evolved, achieves its successes by incorporating and using existing information.

Mechanisms are never self-explanatory. For instance, your Chevy Impala may be the principal mechanism by which you travel to and from work. Yet explaining how that mechanism gets you from home to work and back again does not explain the information required to build it. Likewise, if natural selection, as operating in conjunction with replication, mutation, and other sources of variation, constitutes the primary mechanism responsible for the evolution of life, the information required to originate this mechanism must still be explained. Moreover, by the Law of Conservation of Information, that information cannot be less than the mechanism gives out in searching for and successfully finding biological form and function.

It follows that Dawkins’s characterization of evolution as a mechanism for building up complexity from simplicity fails. For Dawkins, proper scientific explanation is “hierarchically reductionistic,” by which he means that “a complex entity at any particular level in the hierarchy of organization” must be explained “in terms of entities only one level down the hierarchy.”⁷⁵ Thus, according to Dawkins, “the one thing that makes evolution such a neat theory is that it explains how organized complexity can arise out of primeval simplicity.”⁷⁶ This is also why Dawkins regards intelligent design as unacceptable:

To explain the origin of the DNA/protein machine by invoking a supernatural [*sic*] Designer is to explain precisely nothing, for it leaves unexplained the origin of the Designer. You have to say something like “God was always there,” and if you allow yourself that kind of lazy way out, you might as well just say “DNA was always there,” or “Life was always there,” and be done with it.⁷⁷

Conservation of information shows that Dawkins's primeval simplicity is not as nearly simple as he makes out. Indeed, what Dawkins regards as intelligent design's predicament of failing to explain complexity in terms of simplicity now confronts materialist theories of evolution as well. In *Climbing Mount Improbable*, Dawkins argues that biological structures that at first blush seem vastly improbable with respect to a blind search become quite probable once the appropriate evolutionary mechanism is factored in to revise the probabilities.⁷⁸ But this revision of probabilities just means that a null search has given way to an alternative search. And the information that enables the alternative search to be successful now needs itself to be explained. Moreover, by the Law of Conservation of Information, that information is no less than the information that the evolutionary mechanism exhibits in outperforming blind search. The preceding block quote, which was intended as a refutation of intelligent design, could therefore, with small modifications, be turned against natural selection as well:

To explain the origin of the DNA/protein machine by invoking *natural selection* is to explain precisely nothing, for it leaves unexplained the origin of *the information that natural selection requires to execute evolutionary searches*. You have to say something like "*the information* was always there," and if you allow yourself that kind of lazy way out, you might as well just say "DNA was always there," or "Life was always there," and be done with it.⁷⁹

Conservation of information therefore points to an information source behind evolution that imparts at least as much information to the evolutionary process as this process in turn is capable of expressing by producing biological form and function. As a consequence, such an information source has three remarkable properties: (1) it cannot be reduced to purely material or natural causes; (2) it shows that we live in an informationally porous universe; and (3) it may rightly be regarded as intelligent. The Law of Conservation of Information therefore counts as a positive reason to accept intelligent design. In particular, it establishes ID's scientific bona fides.

Several attempts have been made to block NFL- and LCI-inspired design inferences as they apply to biology. Olle Häggström, a Swedish mathematician, thinks that the no free lunch theorems have been, as he puts it, "hype[d]." According to him, "any realistic model for a fitness landscape will have to exhibit a considerable amount of what" he calls "*clustering*, meaning that similar DNA sequences will tend to produce similar fitness values much more often than could be expected under [a model that allows all possible fitness landscapes]."⁸⁰ For Häggström, "realistic models" of fitness presuppose "geographical structures," "link structures," search space "clustering," and smooth surfaces conducive to "hill climbing."⁸¹ All such structures, however, merely reinforce the teleological conclusion we are drawing, which is that the success of evolutionary search depends on the front-loading or environmental contribution of active information. Simply put, if a realistic model of evolutionary processes employs less than the full complement of fitness functions, that's because active information was employed to constrain their permissible range.

Building on Häggström's criticism, statistician Ronald Meester also questions the applicability of NFL (and by implication LCI) to biology. Like Häggström, Meester sees the NFL theorems as dealing "with an extreme situation: averaging over all fitness functions."⁸² A less "extreme," or more "realistic," model would focus on fewer fitness functions. But, as already noted, any such model, by focusing on fewer fitness functions, needs to constrain the total space of fitness functions, and any such constraint entails an imposition of active information. Yet Meester also extends Häggström's argument:

In a search algorithm as defined in the present article, a sufficiently special target set can

only be reached when the search algorithm is very carefully tailored around the fitness function. This conclusion is a direct consequence of our discussion of the NFL theorems and beyond. This implies that this special target can only be reached by programming with insight into the future. Since Darwinian evolution cannot look into the future, this forces us to conclude that simulations cannot be used for the purpose of explaining how complex features arise into the universe.⁸³

But how does Meester know that Darwinian evolution cannot look into the future? Certainly it is part of the popular mythology associated with Darwinism that it is a nonteleological theory. We quoted Jerry Coyne to that effect in section 1. Such quotes appear across the Darwinian literature. But how do we know that evolution is nonteleological or that any teleology in it must be scientifically unascertainable. Imagine you are on an ancient ship and observe a steersman at the helm. The ship traverses difficult waters and reaches port. You conclude that the vessel's trajectory at sea was teleological. Why? Two things: you see a steersman controlling the ship's rudder who, on independent grounds, you know to be a teleological agent; also, you witness the goal-directed behavior of the ship in finding its way home.

Now imagine a variation on this story. An ancient sailor comes on board a twenty-first century ship that is completely automated so that a computer directly controls the rudder and guides the vessel to port. No humans are on board other than this sailor. Being technologically challenged, he will have no direct evidence of a teleological agent guiding the ship—no steersman of the sort that he is used to will be evident. And yet, by seeing the ship traverse difficult channels and find its way home by exactly the same routes he took with ancient ships guided by human steersmen, he will be in his rights to conclude that a purpose is guiding the ship even if he cannot uncover direct empirical evidence of an embodied teleological agent at the helm.⁸⁴

Now, the Law of Conservation of Information gives this conclusion extra quantitative teeth. According to LCI, any search process that exhibits information by successfully locating a target must have been programmed with no less than what we defined as the active information. Thus, armed with LCI, our ancient steersman, however technologically challenged otherwise, could reasonably infer that a teleological agent had put the necessary active information into the ship (the ship, after all, is not eternal and thus its information could not have resided in it forever). Like the ancient sailor, we are not in a position to, as it were, open the hood of the universe and see precisely how the information that runs evolution was programmed (any more than the sailor can peer into the ship's computers and see how it was programmed). But LCI guarantees that the programming that inserts the necessary information is nonetheless there in both instances.

Stricter Darwinists may resist this conclusion, but consider the alternative. When we run search algorithms in evolutionary computing, we find that these searches are inherently teleological (Meester readily concedes this point). So, we may ask, do such mathematical models adequately represent biological evolution? In these models, careful tailoring of fitness functions that assist in locating targets is always present and clearly teleological. If these models adequately represent biological evolution, then this teleological feature of fitness ought to be preserved in nature, implying that Darwinian evolution is itself teleological.

To avoid this conclusion, we must therefore hold that these models somehow fail to capture something fundamental about the inherent nonteleological character of nature. But on what basis? It's only in virtue of such models that we can be said to have a scientific theory of evolution at all. But no nonteleological mathematical models of Darwinian evolution are known. All of them readily submit to the Law of Conservation of Information. Thus, to deny that these models adequately represent biological evolution is to deny that we have an adequate model of

evolution at all. But in that case, we have no scientific basis for rejecting teleology in evolution. Without a clear mathematical underpinning, evolution degenerates into speculation and mystery-mongering.

Meester claims that these models are irrelevant to biology because Darwinian evolution is inherently nonteleological. But he simply begs the question. Darwinian evolution, as it plays out in real life, could potentially look into the future (and thus be teleological) if the fitness it employed were, as Meester puts it, “programmed with insight into the future.” And how do we know that it isn’t? The search algorithms in evolutionary computing give rampant evidence of teleology—from their construction to their execution to the very problems they solve. So too, when we turn to evolutionary biology, we find clear evidence of teleology: despite Dawkins’s denials, biological evolution is locating targets. Indeed, function and viability determine evolution’s targets (recall section 3) and evolution seems to be doing a terrific job finding them. Moreover, given that Darwinian evolution is able to locate such targets, LCI underwrites the conclusion that Darwinian evolution is teleologically programmed with active information.

Häggström and Meester, in mistakenly criticizing the applicability of NFL- and LCI-inspired design inferences to biology, are at least asking the right questions. Biologists Allen Orr and Richard Dawkins, in criticizing such inferences, display conceptual and technical confusion. Orr’s criticism centers on the capacity of fitness to change as populations evolve. Fitness coevolves with an evolving population, and this, for Orr, allows natural selection to work wonders where otherwise it might be hampered. In reviewing William Dembski’s *No Free Lunch*, Orr writes,

Consider fitness functions that are as unsmooth as you like, i.e., rugged ones, having lots of peaks and few long paths up high hills. (These are the best studied of all fitness landscapes.) Now drop many geographically separate populations on these landscapes and let them evolve independently. Each will quickly get stuck atop a nearby peak. You might think then that Dembski’s right; we don’t get much that’s interesting. But now change the environment. This shifts the landscape’s topography: a sequence’s fitness isn’t cast in stone but depends on the environment it finds itself in. Each population may now find it’s no longer at the best sequence and so can evolve somewhat even if the new landscape is still rugged. Different populations will go to different sequences as they live in different environments. Now repeat this for 3.5 billion years. Will this process yield interesting products? Will we get different looking beasts, living different kinds of lives? My guess is yes.⁸⁵

Guessing aside, a straightforward mathematical analysis settles the matter.⁸⁶ It is quite true that fitness can change over time. But nothing in our mathematical development of LCI requires static, time-independent fitness functions. Given an m -query search of a space Ω , it can be represented as a 1-query search of the m -fold Cartesian product Ω^m . Fitness on this product may well change from factor to factor. But this observation poses no challenge to LCI. Mathematics is well able to accommodate Orr’s coevolving fitness functions. Coevolving fitness functions leave the Law of Conservation of Information intact. Note that Wolpert and Macready, when they first stated the NFL theorems, proved them for time-independent as well as for time-dependent fitness functions.⁸⁷ Both NFL and LCI apply to coevolving fitness functions.

Finally, we consider a criticism by Richard Dawkins. According to it, design inferences in biology are rendered superfluous because of the presumed ability of selection to cumulate biological function and complexity by small incremental steps (each of which is quite probable). Dawkins’s WEASEL is his best known illustration of cumulative selection. We analyzed the

WEASEL program in sections 3 and 5, showing how active information was inserted into it to ensure that it located the target phrase METHINKS•IT•IS•LIKE•A•WEASEL. Dawkins has asserted the cumulative power of natural selection for over twenty years, going back to his 1986 book *The Blind Watchmaker*, repeating the claim in his 1996 book *Climbing Mount Improbable*, and repeating it yet again in his most recent 2006 book *The God Delusion*, where he writes,

What is it that makes natural selection succeed as a solution to the problem of improbability, where chance and design both fail at the starting gate? The answer is that natural selection is a cumulative process, which breaks the improbability up into small pieces. Each of the small pieces is slightly improbable, but not prohibitively so.⁸⁸

This claim cannot be sustained in the face of LCI. Dawkins here describes an evolutionary process of, let us say, m steps (or “pieces”), each of which is sufficiently probable that it could reasonably happen by chance. But even m highly probable events, if occurring independently, can have a joint probability that’s extremely low. Dawkins himself makes this point: “When large numbers of these slightly improbable events are stacked up in series, the end product of the accumulation is very very improbable indeed, improbable enough to be far beyond the reach of chance.”⁸⁹ Dawkins here tacitly presupposes an evolutionary search space that consists of an m -fold Cartesian product. Moreover, he rightly notes that uniform probability on this space (which is the product of uniform probabilities on the individual factors—this is elementary probability theory) assigns a very low probability (high endogenous information) to such evolutionary events.

In consequence, for natural selection to be a powerful cumulative force, it’s not enough that the individual steps in Dawkins’s evolutionary process be reasonably probable (or only “slightly improbable”); rather, all these reasonably probable steps must, *when considered jointly*, also be reasonably probable. And this just means the Dawkins has, with zero justification, substituted an alternative search for a null search. But whence this alternative search? Darwinian theory has no answer to this question. To paraphrase Dawkins, “You have to say something like ‘the alternative search just magically appeared or was always there,’ and if you allow yourself that kind of lazy way out, you might as well just say ‘DNA was always there,’ or ‘Life was always there,’ and be done with it.”⁹⁰ Cumulative selection, as Dawkins characterizes it, does nothing to explain the source of evolution’s creative potential. LCI, by contrast, shows that evolution’s creative potential lies in its incorporation and use of active information.

9 Conclusion: “A Plan for Experimental Verification”

In a 2002 address to the American Scientific Affiliation, Francis Collins, then director of the NIH’s National Human Genome Research Institute, posed the following challenge to intelligent design: “A major problem with the Intelligent Design theory is its lack of a plan for experimental verification.”⁹¹ We submit that the Law of Conservation of Information makes such a plan feasible.

The Law of Conservation of Information states that active information, like money or energy, is a commodity that obeys strict accounting principles. Just as corporations require money to power their enterprises and machines require energy to power their motions, so searches require active information to power their success. Moreover, just as corporations need to balance their books and machines cannot output more energy than they take in, so searches, in successfully locating a target, cannot give out more information than they take in.

It follows from the Law of Conservation of Information that active information cannot be

gotten on the cheap but must always be paid for in kind. As such, this law has far reaching implications for evolutionary theory, pointing up that the success of evolutionary processes in exploring biological search spaces always depends on preexisting active information. In particular, evolutionary processes cannot create from scratch the active information that they require for successful search.

The Law of Conservation of Information, however, is not merely an accounting tool. Under its aegis, intelligent design merges theories of evolution and information, thereby wedding the natural, engineering, and mathematical sciences. On this view (and there are other views of intelligent design), its main focus becomes how evolving systems incorporate, transform, and export information.⁹² Moreover, a principal theme of its research becomes teasing apart the respective roles of internally produced and externally applied information in the performance of evolving systems.

Evolving systems require active information. How much? Where do they get it? And what does this information enable them to accomplish? Tracking and measuring active information in line with the Law of Conservation of Information is the plan we propose for experimentally verifying intelligent design and thus meeting Collins's challenge. Evolution is a theory of process. It connects dots by telling us that one thing changes into another and then specifies the resources by which the transformation is (supposed to be) effected. According to the Law of Conservation of Information, active information will always constitute a crucial resource in driving targeted evolutionary change (and much of that change in biology, we have argued, is indeed targeted—exceptions include cyclic variation and genetic drift).

Tracking and measuring active information to verify intelligent design is readily achieved experimentally. Consider, for instance, that whenever origin-of-life researchers use chemicals from a chemical supply house, they take for granted information-intensive processes that isolate and purify chemicals. These processes typically have no analogue in realistic prebiotic conditions. Moreover, the amount of information these processes (implemented by smart chemists) impart to the chemicals can be calculated. This is especially true for polymers, whose sequential arrangement of certain molecular bases parallels the coded information that is the focus of Shannon's theory of communication.⁹³ In such experiments, a target invariably exists (e.g., a simple self-replicating molecule, an autocatalytic set, or a lipid membrane).⁹⁴

Just as information needs to be imparted to a golf ball to land it in a hole, so information needs to be imparted to chemicals to render them useful in origin-of-life research. This information can be tracked and measured. Insofar as it obeys the Law of Conservation of Information, it confirms intelligent design, showing that the information problem either intensifies as we track material causes back in time or terminates in an intelligent information source. Insofar as this information seems to be created for free, LCI calls for closer scrutiny of just where the information that was given out was in fact put in.

In such information-tracking experiments, the opponent of intelligent design hopes to discover a free lunch. The proponent of intelligent design, by contrast, attempts to track down hidden information costs and thereby confirm that the Law of Conservation of Information was preserved. There is no great mystery in any of this. Nor do such experiments to confirm intelligent design merely apply to the origin of life. Insofar as evolution (whether chemical or biological) is an exact experimental science, it will exhibit certain informational properties. Are those properties more akin to alchemy, where more information comes out than was put in? Or are they more akin to accounting, where no more information comes out than was put in? A systematic attempt to resolve such questions constitutes a plan for experimentally verifying intelligent design.

Notes

¹Robert Stalnaker, *Inquiry* (Cambridge, Mass.: MIT Press, 1984), p. 85.

²Fred Dretske, *Knowledge and the Flow of Information* (Cambridge, Mass.: MIT Press, 1981), p. 4.

³Douglas Robertson, "Algorithmic Information theory, Free Will, and the Turing Test," *Complexity* 4(3), 1999: 25–34.

⁴G. K. Chesterton, *Orthodoxy*, in *Collected Works of G. K. Chesterton*, vol. 1 (San Francisco: Ignatius, 1986), 243.

⁵Jerry Coyne, "Truckling to the Faithful: A Spoonful of Jesus Makes Darwin Go Down," posted on his blog Why Evolution Is True on April 22, 2009 at <http://whyevolutionistrue.wordpress.com> (last accessed April 27, 2009). Emphasis added. "Genetic drift" here refers to random changes in population gene frequencies. It too is nonteleological.

⁶Darwinist philosopher of biology David Hull explicitly confirms this point: "He [Darwin] dismissed it [design] not because it was an incorrect scientific explanation, but because it was not a proper scientific explanation at all." David Hull, *Darwin and His Critics: The Reception of Darwin's Theory of Evolution by the Scientific Community* (Cambridge, Mass.: Harvard University Press, 1973), 26.

⁷See, for instance, F. H. Sandbach, *The Stoics*, 2nd ed. (Indianapolis: Hackett, 1989), 72–75.

⁸Aristotle, *Metaphysics*, trans. W. D. Ross, XII.3 (1070a, 5–10), in Richard McKeon, ed., *The Basic Works of Aristotle* (New York: Random House, 1941), 874.

⁹Aristotle, *Physics*, trans. R. P. Hardie and R. K. Gaye, II.8 (199a, 15–20), in Richard McKeon, ed., *The Basic Works of Aristotle* (New York: Random House, 1941), 250.

¹⁰*Ibid.*, II.8 (199b, 25–30), 251.

¹¹Compare Jacques Monod, *Chance and Necessity* (New York: Vintage, 1972).

¹²David Baltimore, "DNA Is a Reality beyond Metaphor," *Caltech and the Human Genome Project* (2000): available online at <http://pr.caltech.edu:16080/events/dna/dnabalt2.html> (last accessed April 23, 2007).

¹³Manfred Eigen, *Steps Towards Life: A Perspective on Evolution*, trans. Paul Woolley (Oxford: Oxford University Press, 1992), 12.

¹⁴Eörs Szathmáry and John Maynard Smith, "The Major Evolutionary Transitions," *Nature* 374 (1995): 227–232.

¹⁵Holmes Rolston III, *Genes, Genesis and God: Values and Their Origins in Natural and Human History* (Cambridge: Cambridge University Press, 1999), 352.

¹⁶"Give me twenty-six lead soldiers and I will conquer the world." The "lead soldiers" are the typefaces used in printing. This quote has variously been attributed to Karl Marx, William Caxton, and Benjamin Franklin. Reliable references appear to be lacking.

¹⁷Francisco J. Ayala, "Darwin's Revolution," in *Creative Evolution?!*, eds. J. H. Campbell and J. W. Schopf (Boston: Jones and Bartlett, 1994), 4. The subsection from which this quote is taken is titled "Darwin's Discovery: Design without Designer."

¹⁸Richard Dawkins, *The Blind Watchmaker: Why the Evidence of Evolution Reveals a Universe without Design* (New York: Norton, 1987), 6.

¹⁹Francis S. Collins, *The Language of God: A Scientist Presents Evidence for Belief* (New York: Free Press, 2006), 201.

²⁰Theistic evolutionist Kenneth Miller justifies the scientific undetectability of divine action as follows: "The indeterminate nature of quantum events would allow a clever and subtle God to influence events in ways that are profound, but *scientifically undetectable* to us. Those events could include the appearance of mutations, the activation of individual neurons in the brain, and even the survival of individual cells and organisms affected by the chance processes of radioactive decay." Kenneth R. Miller, *Finding Darwin's God: A Scientist's Search for Common Ground between God and Evolution* (New York: Harper, 1999), 241. Emphasis added.

²¹Richard Dawkins, *The Blind Watchmaker* (New York: Norton, 1986), pp. 47–48.

²²For an event of probability p to occur at least once in k independent trials has probability $1 - (1-p)^k$ —see Geoffrey Grimmett and David Stirzaker, *Probability and Random Processes* (Oxford: Clarendon, 1982), 38. If p is small and $k = 1/p$, then this probability is greater than $1/2$. But if k is much smaller than $1/p$, this probability will be quite small (i.e., close to 0).

²³Dawkins, *Blind Watchmaker*, p. 48.

²⁴Richard Dawkins, *The Blind Watchmaker* (New York: Norton, 1987), 50.

²⁵See <http://www.top500.org/lists/2008/11> (last accessed April 20, 2009).

²⁶“Workers at the University of Georgia estimate that 10^{30} single-celled organisms are produced every year; over the billion-year-plus history of the earth, the total number of cells that have existed may be close to 10^{40} .” Michael J. Behe, *The Edge of Evolution: The Search for the Limits of Darwinism* (New York: Free Press, 2007), 153.

²⁷Seth Lloyd, “Computational Capacity of the Universe,” *Physical Review Letters* 88(23) (2002): 7901–7904.

²⁸J. Bowie and R. Sauer, “Identifying Determinants of Folding and Activity for a Protein of Unknown Sequences: Tolerance to Amino Acid Substitution,” *Proceedings of the National Academy of Sciences* 86 (1989): 2152–56. J. Bowie, J. Reidhaar-Olson, W. Lim, and R. Sauer, “Deciphering the Message in Protein Sequences: Tolerance to Amino Acid Substitution,” *Science* 247 (1990):1306–10. J. Reidhaar-Olson and R. Sauer, “Functionally Acceptable Solutions in Two Alpha-Helical Regions of Lambda Repressor,” *Proteins, Structure, Function, and Genetics* 7 (1990): 306–10. See also Michael Behe, “Experimental Support for Regarding Functional Classes of Proteins to be Highly Isolated from Each Other,” in *Darwinism: Science or Philosophy?*, eds. J. Buell, and G. Hearn (Dallas: Foundation for Thought and Ethics, 1994), 60–71; and Hubert Yockey, *Information Theory and Molecular Biology* (Cambridge: Cambridge University Press, 1992), 246–58.

²⁹Nils Aall Barricelli, “Numerical Testing of Evolution Theories, Part I: Theoretical Introduction and Basic Tests,” *Acta Biotheoretica* 16(1–2) (1962): 69–98. Reprinted in David B. Fogel, ed., *Evolutionary Computation: The Fossil Record* (Piscataway, N.J.: IEEE Press, 1998), 166.

³⁰J. L. Crosby, “Computers in the Study of Evolution,” *Science Progress, Oxford* 55 (1967): 279–292. Reprinted in Fogel, *Evolutionary Computation*, 95.

³¹Heinz R. Pagels, *The Dreams of Reason: The Computer and the Rise of Sciences of Complexity* (New York: Simon and Schuster, 1989), 104.

³²Robert T. Pennock, “DNA by Design?” in W. A. Dembski and M. Ruse, eds., *Debating Design: From Darwin to DNA* (Cambridge: Cambridge University Press, 2004), 141. Pennock’s work on AVIDA is summarized in Richard Lenski, Charles Ofria, Robert T. Pennock, and Christoph Adami, “The Evolutionary Origin of Complex Features,” *Nature* 423 (May 8, 2003): 139–144.

³³MESA, which stands for Monotonic Evolutionary Simulation Algorithm, is available at <http://www.iscid.org/mesa> (last accessed April 6, 2009).

³⁴See <http://www.i2cam.org> (last accessed April 6, 2009).

³⁵Richard E. Lenski, “Phenotypic and Genomic Evolution during a 20,000-Generation Experiment with the Bacterium *Escherichia coli*,” *Plant Breeding Reviews* 24 (2004): 225–265.

³⁶Michael Behe, *Darwin’s Black Box: The Biochemical Challenge to Evolution* (New York: Free Press, 1996).

³⁷Lenski et al., “The Evolutionary Origin of Complex Features.”

³⁸Quoted in John Horgan, “From Complexity to Perplexity: xxx,” *Scientific American*

³⁹David Berlinski, *The Devil’s Delusion: Atheism and Its Scientific Pretensions* (New York: Crown Forum, 2008), 190.

⁴⁰For Adami’s AVIDA, visit his Digital Life Lab at Caltech at <http://dllib.caltech.edu>. For Ray’s Tierra, go to <http://life.ou.edu/tierra>. For Schneider’s *ev*, go to <http://www-lmmb.ncifcrf.gov/~toms/paper/ev>. Each of these sites was last accessed April 6, 2009. See also Thomas D. Schneider, “Evolution of Biological Information,” *Nucleic Acids Research* 28(14) (2000): 2794–2799.

⁴¹These computer programs that suggest sharp limits to the power of the Darwinian selection mechanism. For Mendel's Accountant visit <http://mendelsaccount.sourceforge.net> and for MutationWorks visit <http://www.mutationworks.com>. Both these sites were last accessed April 6, 2009.

⁴²The alternative is for evolution to constitute an inexact historical science: "To obtain its answers, particularly in cases in which experiments are inappropriate, evolutionary biology has developed its own methodology, that of *historical narratives* (tentative scenarios)." Quoted from Ernst Mayr, *What Makes Biology Unique? Considerations on the Autonomy of a Scientific Discipline* (Cambridge: Cambridge University Press, 2004), 24–25. Emphasis in the original.

⁴³Kenneth R. Miller, *Only a Theory: Evolution and the Battle for America's Soul* (New York: Viking, 2008), 77–78.

⁴⁴See the Evolutionary Informatic Lab's "EV Ware: Dissection of a Digital Organism" at <http://www.evoinfo.org/Resources/EvWare/index.html> (last accessed April 6, 2009).

⁴⁵For the details see William A. Dembski, *No Free Lunch: Why Specified Complexity Cannot Be Purchased without Intelligence* (Lanham, Md.: Rowman and Littlefield, 2002), sec. 4.9.

⁴⁶Leon Brillouin, *Science and Information Theory*, 2nd ed. (New York: Academic Press, 1962), 269.

⁴⁷Peter Medawar, *The Limits of Science* (Oxford: Oxford University Press, 1984), 79.

⁴⁸William A. Dembski and Robert J. Marks II, "The Search for a Search: Measuring the Information Cost of Higher Level Search," [redacted], forthcoming. William A. Dembski and Robert J. Marks II, "The Conservation of Information: Measuring the Information Cost of Successful Search," [redacted], forthcoming. For these and additional papers on tracking the active information in intelligent computing see the publications page at <http://www.evoinfo.org>.

⁴⁹Actually, the problem of identifying an appropriate fitness function for locating Dawkins's target phrase is much worse than sketched here. We focused entirely on the subset of single-hill fitness functions, which could have its peak at any of 10^{40} places. But there are many more fitness functions than those gauging distance from a target sequence. Indeed, the total space of fitness functions is exponential in the underlying sequence space (see the fitness-theoretic version of the conservation of information theorem in section 6) and, properly speaking, would have to be searched. In limiting ourselves to fitness functions based on the Hamming distance from a target sequence, we've already incurred a heavy informational cost.

⁵⁰Jakob Bernoulli, *Ars Conjectandi* (1713; reprinted Ann Arbor, Mich.: University of Michigan Library, 2006). For the applicability of Bernoulli's principle of insufficient reason in the information-theoretic context, see William A. Dembski and Robert J. Marks II, "Bernoulli's *Principle of Insufficient Reason* and Conservation of Information in Computer Search," available at <http://www.evoinfo.org>.

⁵¹Steven Pinker, *How the Mind Works* (New York: Norton, 1999), 90–92.

⁵²This result is proved in Appendix C of Dembski and Marks, "The Search for a Search." A proof is also available online in William A. Dembski, "Searching Large Spaces: Displacement and the No Free Lunch Regress," (2005): http://www.designinference.com/documents/2005.03.Searching_Large_Spaces.pdf.

⁵³See Robert J. Marks II, *Handbook of Fourier Analysis and Its Applications* (Oxford: Oxford University Press, 2009), 165.

⁵⁴Joseph C. Culberson "On the Futility of Blind Search: An Algorithmic View of 'No Free Lunch'," *Evolutionary Computation* 6(2) (1998): 109–127.

⁵⁵Christian Igel and Marc Toussaint, "On Classes of Functions for Which No Free Lunch Results Hold," *Information Processing Letters* 86 (2003): 317–321. An earlier version of this paper is available online at <http://www.marc-toussaint.net/publications/igel-toussaint-01.pdf> (last accessed April 13, 2009).

⁵⁶See the information bounds that come up in the Vertical No Free Lunch Theorem that is proved in Dembski and Marks, "The Search for a Search."

⁵⁷David H. Wolpert and William G. Macready, "No Free Lunch Theorems for Optimization," *IEEE Transactions on Evolutionary Computation* 1(1) (1997): 67–82.

⁵⁸Cullen Schaffer, “A Conservation Law for Generalization Performance,” *Machine Learning: Proceedings of the Eleventh International Conference*, eds. H. Hirsh and W. W. Cohen, 259–265 (San Francisco: Morgan Kaufmann, 1994).

⁵⁹Cullen Schaffer, “Conservation of Generalization: A Case Study,” typescript (1995): available online at <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.41.672> (last accessed April 23, 2009). Emphasis in the original.

⁶⁰Yu-Chi Ho and David L. Pepyne, “Simple Explanation of the No Free Lunch Theorem,” *Proceedings of the 40th IEEE Conference on Decision and Control*, Orlando, Florida (2001): 4409–4414.

⁶¹For the Church-Turing Thesis, often simply called Church’s Thesis, see Klaus Weihrauch, *Computability* (Berlin: Springer-Verlag, 1987), sec. 1.7.

⁶²Keith Devlin, *Logic and Information* (Cambridge: Cambridge University Press, 1991), 2.

⁶³“Maxwell’s demon inhabits a divided box and operates a small door connecting the two chambers of the box. When he sees a fast molecule heading toward the door from the far side, he opens the door and lets it into his side. When he sees a slow molecule heading toward the door from his side he lets it through. He keeps the slow molecules from entering his side and the fast molecules from leaving his side. Soon, the gas in his side is made up of fast molecules. It is hot, while the gas on the other side is made up of slow molecules and it is cool. Maxwell’s demon makes heat flow from the cool chamber to the hot chamber.” John R. Pierce, *An Introduction to Information Theory: Symbols, Signals and Noise*, 2nd rev. ed. (New York: Dover, 1980), 199.

⁶⁴Leo Szilard, “Über die Entropieverminderung in einem thermodynamischen System bei Eingriff intelligenter Wesen,” *Zeitschrift für Physik* 53 (1929): 840–856. For the English translation see John A. Wheeler and Wojciech H. Zurek, eds., *Quantum Theory and Measurement* (Princeton: Princeton University Press, 1983), 539–548.

⁶⁵John Stuart Mill, *A System of Logic: Ratiocinative and Inductive*, 8th ed. (1882; reprinted London: Longmans, Green, and Co., 1906).

⁶⁶*Ibid.*, 256.

⁶⁷Miller, *Only a Theory*, 77.

⁶⁸Barricelli, “Numerical Testing of Evolution,” 170–171. Emphasis added to underscore that Barricelli cites exactly the same three evolutionary mechanisms as Miller: selection, replication (or reproduction), and mutation.

⁶⁹*Ibid.*, see especially secs. 3–6.

⁷⁰See Lynn Margulis and Dorion Sagan, *Acquiring Genomes: A Theory of the Origins of Species* (New York: Basic Books, 2002). Margulis’s work on symbiogenesis goes back now over thirty years.

⁷¹Quoted in Michael Shermer, “The Woodstock of Evolution,” *Scientific American* (June 27, 2005): available online at <http://www.sciam.com/article.cfm?id=the-woodstock-of-evolutio&print=true> (last accessed April 27, 2009). Emphasis added. The occasion for Margulis’s remark was her award of an honorary doctorate at the World Summit on Evolution, Galapagos Islands, June 8 – 12, 2005.

⁷²Brian Goodwin, *How the Leopard Changed Its Spots: The Evolution of Complexity* (New York: Scribner’s, 1994), 35–36.

⁷³Stuart Kauffman, *Investigations* (New York: Oxford University Press, 2000), 19.

⁷⁴*Ibid.*, 18.

⁷⁵Dawkins, *Blind Watchmaker*, 13.

⁷⁶*Ibid.*, 316.

⁷⁷*Ibid.*, 141. Note that intelligent design is not committed to supernaturalism. Stoic and Aristotelian philosophy eschewed the supernatural but were clearly in the ID camp.

⁷⁸Richard Dawkins, *Climbing Mount Improbable* (New York: Norton, 1996).

⁷⁹For the original, unedited passage, see Dawkins, *Blind Watchmaker*, 141. The words in italics substitute for words in the original. The words in italics are about natural selection and information; the words in the original that they replace are about God and design.

⁸⁰Olle Häggström, “Intelligent Design and the NFL Theorem,” *Biology and Philosophy* 22 (2007): 226, 228.

⁸¹*Ibid.*, 220, 228.

⁸²Ronald Meester, “Simulation of Biological Evolution and the NFL Theorems,” *Biology and Philosophy*, forthcoming.

⁸³*Ibid.*

⁸⁴Our ancient sailor is, we might say, undergoing a Turing Test in which he must distinguish the unmediated teleology in a ship’s course due to the direct activity of an embodied teleological agent versus the mediated teleology in a ship’s course due to its having been programmed by a teleological agent. Without actually seeing the steersman, the sailor would be unable to distinguish the two forms of teleology. Such indistinguishability, according to Alan Turing, would suggest that the teleology is as real in the one case as in the other. For the Turing Test, see Alan Turing, “Computing Machinery and Intelligence,” *Mind* 59 (1950): 434–460.

⁸⁵H. Allen Orr, “Book Review: *No Free Lunch*,” *Boston Review* (Summer 2002): available online at <http://bostonreview.net/BR27.3/orr.html> (last accessed April 28, 2009). Compare Dembski’s response to Orr, “Evolution’s Logic of Credulity,” (2002): available online at http://www.designinference.com/documents/2002.12.Unfettered_Resp_to_Orr.htm (last accessed April 28, 2009).

⁸⁶Such an analysis was offered in section 4.10 of Dembski’s *No Free Lunch*, a section Orr’s review failed to address.

⁸⁷Wolpert and Macready, “No Free Lunch Theorems for Optimization.” They prove an NFL theorem for “static cost functions” (i.e., time-independent fitness functions) as well as for “time-dependent cost functions” (i.e., time-dependent fitness functions).

⁸⁸Richard Dawkins, *The God Delusion* (London: Bantam Press, 2006), 121.

⁸⁹*Ibid.* For instance, it’s highly probable that any particular person will avoid death in an automobile accident today. But given that the number of drivers m is large enough, it becomes highly improbable that none of these drivers will be killed.

⁹⁰For the original, unedited passage, see Dawkins, *Blind Watchmaker*, 141.

⁹¹Francis S. Collins, “Faith and the Human Genome,” *Perspectives on Science and Christian Faith* 55(3) (2003): 151.

⁹²The form of intelligent design that we are describing here falls under what we have dubbed “evolutionary informatics”—see <http://www.evoinfo.org>.

⁹³Claude Shannon and Warren Weaver, *The Mathematical Theory of Communication* (Urbana, Ill.: University of Illinois Press, 1949).

⁹⁴See William A. Dembski and Jonathan Wells, *How to Be an Intellectually Fulfilled Atheist (Or Not)* (Wilmington, Del.: ISI Books, 2008), ch. 19, titled “The Medium and the Message.” This book critically analyzes origin-of-life theories.