Information and the Argument from Design

(Online Version)

Peter Godfrey-Smith

Philosophy Program, RSSS
Australian National University
and
Philosophy Department
Harvard University

1. Introduction

William Dembski holds that "the origin of information is best sought in intelligent causes" ("Intelligent Design as a Theory of Information", 1997). In particular, Dembski argues that Darwinism is not able to explain the existence of biological structures that contain a certain kind of information – "complex specified information" (CSI). To explain these informational features of living systems, we must instead appeal to the choices made by an intelligent designer.

Dembski's version of the argument from design derives an apparent sophistication from the information-theoretic terms in which it is expressed. But in the end, Dembski's version is one of the least plausible versions of the design argument. Partly because of the formal apparatus, Dembski's version of the argument is far too sweeping; it omits qualifications that other opponents of Darwinism sensibly include.

My discussion will have two main sections. One outlines information theory, and Dembski's use of the concepts of information and complexity. The other looks at the alleged consequences of these ideas for Darwinism.[1]

2. Information and Probability

Dembski uses the resources of information theory to describe the phenomena he thinks natural causes cannot explain. At bottom, the only thing information theory does in his argument is provide an alternative way of expressing the idea that some events are *highly unlikely to have arisen by chance*, and unlikely in a way that can be described without hindsight. Events of this kind require a special type of explanation – on that point Dembski and many Darwinians agree. But Dembski argues further that no set of natural causes can provide the required explanation.

Information theory, as Dembski says, starts from the idea of contingency – from the idea of a situation in which one out of a range of possible outcomes is actualized. The concept of information, Dembski says, does not require the existence of a channel transmitting the information, or anyone receiving it. Anything which has a range of possible states or outcomes, where these states or outcomes each have some definite probability of occurring, has information associated with it. The more different possible states a "source" of this kind has, and the more similar are the probabilities of these states, the more information is associated with the source.

Describing information as something that exists at a source independently of any channel or receiver is not completely uncontroversial within information theory.[2] Some prefer to only use the term "entropy" for the property of the source itself, and to introduce the term "information" when the source is being considered in relation to another variable with which it might be associated (see Cover and Thomas 1991 chapter 2). But this is not important to the argument.

The idea that information is associated with uncertainty or contingency is one basic idea in information theory. Another has to do with the transmission of information. If there are two variables (A and B) which both have some uncertainty associated with them, and where the state of variable A is reliably associated with the state of B, then the state of A carries information about the state of B. That is, if the reduction of uncertainty at A also reduces uncertainty at B, then A carries information about B. The number of tree rings in a tree stump, for example, carries information about the age of the tree. So information, in this sense, is something that

exists wherever there is measurable uncertainty. And information is transmitted by means of physical correlations.[3]

Information theory itself is mostly concerned with the mathematical treatment of relationships between quantities of information. The role point of the theory is to describe the transmission of information over channels, the storage of information in memory, and other kinds of use and manipulation of information. A paradigm situation for the theory is one where we have a source of information, a receiver of information about the source, and a channel that connects them and makes it possible for information to flow from source to receiver, in the form of signals. The theory can be used to describe which kinds of channels have the capacity to transmit messages with given mathematical properties. To pick a very simple example, if a source has four possible states, in order for the receiver to be able to tell from a signal which one of these four states is actual, the signal must have at least four possible states. Otherwise the signal must be ambiguous. Although "flow" of information has an evocative sound, all that is required for information flow is a certain kind of reliable association between what happens at the source and the state of a signal received by the receiver.

Applications like Dembski's make use of just one part of this theoretical structure – the first of the basic ideas listed above. This is the idea that a source of information (or "probabilistic source") is anything with a number of distinct possible states, where the states each have a probability of occurring.

The point about probability is important; in order to use information theory in some domain, we have to be able to ascribe an *objective probability* to each possible state of the source. Ascribing objective probabilities to some events, like the outcomes of lotteries, is often regarded as fairly easy. Other events are much harder to describe in this way, not just because the probabilities are well-hidden but because it is hard to know what is even meant by the description. Unique and complex events are hard cases, for example. Ancient Rome was first sacked by Visigoths, in 410 AD. This was a contingent matter; other tribes might have done the job instead, and it might have happened earlier or later. So we have a space of possibilities with only one being actualized. What was the probability that Rome be first sacked by the Celts, or the Chinese, or Native Americans? In some sense these alternatives seem unlikely, but in describing a unique and complex event (as opposed to the outcome of a gambling game which is repeated over and over) it is very hard to ascribe probabilities to specific outcomes, and many statisticians would refuse to even try. To use

information theory, though, you have to have probabilities. So even though we believe that Rome's sacking was a contingent matter, this is apparently not the kind of case where we can say that there is some definite "amount of information" associated with the fact that the Visigoths first sacked Rome in 410 AD. Just because some event can be thought about in contrast with a range of possible alternatives, this does not imply that there is a definite and measurable amount of information associated with that event.

Though I think it is very implausible to claim there is a definite amount of information associated with every contingent event or state of affairs, in this paper I will concede to Dembski all the assumptions he needs about the existence of physical probabilities, and their connection to information. I do this partly for the sake of discussion, but also because many biologists, including Darwinians, are happy to use probabilistic and information-theoretic frameworks in roughly the same way that Dembski does. So it is better to concede these issues to Dembski for now, and reply to him in a way that does not require controversial assumptions about probability.

For a particular event or structure to contain information, there must be a range of alternatives which could have been actual instead. Information theory is usually not much concerned with specific events. But as Dembski says, it is possible in principle to take a single, specific event which has a single probability, and assign a measure of information to that event. The more improbable the event, the more information is associated with it. This measure is sometimes referred to as the "surprisal" of the particular event.

When we are talking only about one specific event, the relationship between a description in terms of probability and a description in terms of information is trivial. To assign a measure of information to the event, you just mathematically transform its probability. You find the logarithm to the base 2 of that probability, and take the negative of that logarithm. A probability of 1/4 becomes 2 bits of information, as the logarithm to the base 2 of 1/4 is -2. A probability of 1/32 becomes 5 bits of information, and so on. In saying these things, we are doing *no more* than applying a mathematical transformation to the probabilities. Because the term "information" is now being used, it might seem that we have done something important. But we have just re-scaled the probabilities that we already had.

So why do the re-scaling at all? Logarithms are convenient to work with for some purposes; if we were going to multiply the probabilities we can add their logarithms instead. But that is just convenience. If we were faced with a problem of

the sort that information theory is intended to deal with, such as the problem of working out whether a particular medium of communication is suitable for the messages we want to send, then there is more we could do. But assessing the power of Darwinism does not have much to do with choosing communication media. Despite all the detail that Dembski gives in describing information theory, information is not making any essential contribution to his argument. What is doing the work is just the idea of objective probability. We have objective probabilities associated with events or states of affairs, and we are re-expressing these probabilities with a mathematical transformation.[4]

So far I have discussed Dembski's use of the term "information." Something should be said also about "complexity" and "specification," as Dembski claims that the problem for Darwinism is found in cases of "complex specified information" (CSI). Do these concepts add anything important to Dembski's argument? "Complexity" as used by Dembski does not add anything, as by "complex information" Dembski just means "lots of information." He uses the distinction between "simple" and "complex" as a way of referring to the amount of information associated with an event.

Dembski motivates his introduction of the loaded term "complex" by linking his discussion to another mathematical framework, "complexity theory." Complexity theory, in the sense Dembski intends, is mostly concerned with the measurement of the complexity of formal problems and computational operations. Computational complexity theory uses a special sense of "complex" which should not be confused with everyday senses of the term, or with biological uses. As in the case of information, Dembski is making his connection to complexity theory in only a minimal way; Dembski says that there is more complexity as the amount of information associated with an event goes up (or equivalently, as the probability of the event goes down). There is nothing stopping Dembski calling this "complexity," but it is important to realize how little the term contributes to the argument.

When a full house is dealt in a hand of poker, we have a more improbable event than we would have had if two pairs had been dealt. So there is more information generated by the full house. And in Dembski's special sense, the full house is associated with "more complex information" than the two pairs – again, he just means the full house is less likely than two pairs. Dembski is saying, in lots of different ways, that some events are less likely to arise by chance than others.

The term "complexity" is ambiguous and hard to define in everyday usage. I doubt that in any everyday sense, a full house is "more complex" than two pairs. And

more importantly, the various biological uses of the term "complex" are not captured simply by the idea of low probability or high informational content. Admittedly it is hard to work out what some biologists mean by the term "complex." I have argued (1996) that the core meaning of complexity, in biological contexts, is *heterogeneity*. Something is complex when it has many different kinds of parts, or does many different things. For me, there are as many kinds of complexity as there are kinds of heterogeneity. John Bonner, in *The Evolution of Complexity* (1988), measures the biological complexity of an organism by counting the number of different types of cell the organism contains. This is one kind of heterogeneity – heterogeneity in cell types – and Bonner's book shows this to be a clear and useful sense of "complexity." Dan McShea (1991) has discussed a number of other senses used by biologists. Some of these are linked to information-theoretic measures, such as the entropy of a probability distribution. But none of the senses of complexity recognized or used by biologists can be defined simply in terms of the improbability of a particular structure or event. In Dembski's article, however, the three phrases "very improbable," "containing a lot of information" and "containing complex information" all pick out the same property.[5]

The final term Dembski introduces is "specified," which also applies to information. Here he is dealing with a genuinely tricky problem. Some events can be truly described as improbable, but only in a sense involving hindsight. If you are dealt a worthless poker hand, it is nonetheless extremely improbable for you to have been dealt *that exact* worthless hand. All poker hands are improbable in that sense, but the game of poker singles out some combinations of cards as valuable, and specifies these ahead of time. So, in many judgments about probability we are interested only in events that are improbable *in a sense that does not involve hindsight*. That is, we assess the probabilities according to some pre-ordained scheme for classifying outcomes (full house, three of a kind, flush, etc.).

That completes the outline of Dembski's information-theoretic framework. Dembski goes on to claim that life contains CSI – complex specified information. This *looks* like an interesting and theoretically rich property, but in fact it is nothing special. Dembski's use of the term "information" should not be taken to suggest that meaning or representation is involved. His use of the term "complex" should not be taken to suggest that things with CSI must be complex in either the everyday sense of the term or a biologist's sense. Anything which is unlikely to have arisen by chance (in a sense which does not involve hindsight) contains CSI, as Dembski has defined it.

So, Dembski's use of information theory provides a roundabout way of talking about probability. Is this probabilistic approach a useful one for tackling general questions about Darwinism? In my own view, there is little to recommend casting these questions in probabilistic terms. They are well enough understood without the probabilistic apparatus, and objective probability is controversial, even in contexts where there is some consensus about the actual judgements. We probably all agree that the chance Rome had of being first sacked by Chinese was lower than its chance of being first sacked by Goths. In saying this, we have in mind the proximity of Rome to the Gothic tribes, the history of antagonism between them, and so on. But it has proved very hard to give a precise meaning to probabilistic claims of this kind. And claims about the probability of life appearing, or of elephants evolving by natural selection, are in the same boat. I think the probabilistic language introduces more problems than it solves here.

In the present context though, it would be unwise for me to reject this way of approaching evolutionary questions. Quite a number of Darwinians see some usefulness in using broad claims about probability, and also claims about information, when discussing Darwinism. G. C. Williams, for example, has placed a lot of emphasis on the role of informational properties in evolution (Williams 1992). Maynard Smith and Szathmáry suggest that various major transitions in evolution resulted from "new ways of transmitting information" across generations (1995 p. 13). Biologists also sometimes use information to discuss the ways in which genes and other developmental factors causally affect biological traits. As DNA sequences make use of a fixed "alphabet" of four elements, the bases C, A T and G, it is possible to discuss the informational content of stretches of DNA. As I said earlier, to associate a measure of information with an event or structure is to view it as one of a definite range of possible alternatives, each of which had some probability of being actual. So each actual base in a DNA sequence is thought of as one of four possible alternatives for that position in the stretch of DNA. If some totally random process was selecting which of the four bases goes in each position, then each actual base in a DNA sequence would be associated with 2 bits of information. The processes that determine the sequence of bases in a particular DNA sequence are in fact far from random, but biologists do sometimes treat a stretch of DNA as containing an amount of information that is a simple function of its length, where each base is treated as in some sense equally likely to be found in each position. A biologist might say: "The information content or determinacy of a complex anatomical structure is orders of

magnitude higher than that of the genome...." (Gilbert, Opitz and Raff 1996 p. 366-367. An organism's genome is its total genetic sequence). What they mean here is that the space of possible genomes (of a certain size) is much smaller than the space of possible anatomical structures that can be relevantly contrasted with the structure in question. The "gap" between the two amounts of information is taken to tell us something about the causation of the anatomical structures. To assert this while using a standard sense of "information content," and without a great deal more probabilistic knowledge, it is necessary to idealize and treat each genome of a certain length as equally likely, and to do the same for each possible anatomical structure. Biologists often do not worry about this idealization; the main aim of the description is just to contrast an actual genetic or phenotypic structure with some range of relevant possible alternatives.

Parts of Dembski's discussion develop in a way that resembles the opening pages of Dawkins' *Blind Watchmaker* (1986). Dawkins says that living things have a special property which makes them very hard to scientifically explain. That property can be described as a certain kind of improbability. And living things are not merely improbable in a sense involving hindsight; Dawkins says the kind of improbability that is relevant here is a kind that must be "specifiable in advance."[6] Both Dembski and Dawkins hold that living things have features that are highly unlikely to have arisen by pure chance, so a special kind of explanation is needed for them.

Whereas the biologists discussed above tend to use information to express the idea that any particular organism is one out of a vast range of possible organisms, and in that sense improbable, Dawkins and Dembski use probability to express the fact that a living organism is one highly organized state for a pile of physical material, material that *could* be combined together in many other ways, most of which would not produce a living system. Life is "improbable" in the sense that there are far more non-living ways for matter to be arranged.

To summarize what we have so far: Dembski uses the idea of information merely to redescribe probabilities. Turning a number into a logarithm does change how you can compute with it, but does not change what (if anything) the number measures. "Information" for Dembski is no more than improbability. "Complex information" is just lots of improbability. An event is associated with "specified" information if the probability associated with it is assigned using a categorization of outcomes which is specifiable in advance and not merely with hindsight.

3. On What Evolutionary Processes Can Bring About

Dembski says that natural causes cannot explain complex specified information. More precisely: an event or structure associated with complex specified information cannot be fully explained in naturalistic terms that do not include some role for intelligent choice. Dembski holds that natural causes cannot create CSI because natural causes are a mixture or combination of "chance" and "necessity." Necessity cannot bring new information into being, and chance can create some information, but not complex specified information. So mutation and natural selection, in particular, cannot create CSI because mutation-plus-selection is just a combination of chance and necessity. Dembski expresses his main claim as a law, the law of the "conservation of information" in all systems of wholly natural causes. Dembski should not really call this a conservation law, because he does allow that CSI can decrease in a system of natural causes. So CSI is not in fact conserved, according to Dembski. But this is a minor issue.

To assess Dembski's claims from a biological standpoint, we need to know what phenomenon is to be explained. Dembski says the phenomenon is "CSI," but where exactly is this CSI found? And if CSI cannot arise from natural causes, which parts of the standard scientific picture of the world – the picture given in standard college biology textbooks, for example – must be changed?

The standard scientific picture includes, firstly, some purely historical claims. The first life appeared on earth over 3.5 billion years ago, as there are fossil traces of bacteria from around that time. There must have been even more basic forms of life prior to bacteria, but these are unknown. For something like 2 billion years, bacteria were the only life on earth, until the appearence of eukaryotic cells, at least 1.7 billion years ago. Multicellular organisms date from about 800 million years ago. Later there appear, among other things, vertebrates (over 500 million years ago), flowering plants (over 120 million years ago), and homo sapiens (about 300,000 years ago). That much is pure history. It is based not just on biological data, but on geology, physics and other sciences.

Along with this historical story we have, in the standard scientific picture, claims about the processes linking these historical events. These events occurred in a system governed, at bottom, only by physical laws – "natural causes" in Dembski's sense. But these physical laws, operating in certain kinds of conditions, give rise to

the processes of replication, variation and natural selection, and all the other factors that contribute to evolution. These factors are taken to be sufficient to generate, on a planet bathed in solar energy and containing the right raw materials, both simple forms of life and then organisms like people arising eventually from organisms like bacteria. So let us ask: if CSI was really "conserved," as Dembski says, which parts of this story would have to be false?

Certainly Dembski claims that the origin of life itself cannot be understood in terms of natural causes, so that part of the standard story must be false. But what about the later events? Dembski also says that "the functionality of organisms" is something that natural causes cannot explain. Dembski does not commit himself to a definition of "functionality" but gestures towards the complicated and hierarchical organization of organisms, their ability to maintain their structure and reproduce, and so on. Here we find CSI.

But this property of "functionality" apparently comes in degrees. Bacteria have a certain amount of "functionality" and a certain amount of CSI, but a camel apparently has a lot more. A bacterium is just a single cell, containing perhaps a few thousand genes. Every cell of the trillions of cells within a camel has a more complex organization than a bacterial cell – though the bacterial cell does have many noteable properties that a camel cell lacks, such as the ability to survive on its own. And a camel contains many different kinds of cell, working together in a hierarchically organized way. It is fairly standard, though not totally uncontroversial, to say that a camel is a more complex organism than a bacterium. Certainly I would say this – in almost all relevant senses, a camel is a more heterogeneous object than a bacterium. So if there is a property of "complex specified information" recognisable in a bacterium, there is apparently more of this same property in a camel.

As a consequence, Dembski seems not to be just denying naturalistic views about the origins of life. He must also hold that getting *from* simpler one-celled organisms *to* more complex, integrated and even intelligent forms of life is not possible with wholly natural causes. So if simpler organisms did appear on earth prior to complex ones, and are connected to them via lines of descent, then the processes by which the more complex came from the less complex must have involved more than natural causes. Intelligent design must have played a role in many later stages, as well as the initial stage, of the history of life.

That is how I interpret Dembski. Possibly he has something else in mind – maybe the earth 3 billion years ago contained both bacteria and a store of hidden CSI which

was slowly incorporated, via natural processes, into sharks, camels and all the rest. As it seems unlikely that camel CSI, orchid CSI and salamander CSI could all be hidden inside bacteria, the extra CSI must have been distributed somehow in the non-biological part of the world before being absorbed into organisms. This is a picturesque option. Or perhaps Dembski thinks that God called everything into being less than 10,000 years ago, and denies that species are linked to each other by any kind of "descent with modification." Dembski presents his argument in abstract terms so it is hard to tell.[7] But it is very important to insist that views like Dembski's confront the *specific historical claims* made by the standard scientific view. These claims concern both the order in which organisms appeared on earth, and the lines of descent that connect them. Either bacteria appeared on earth before camels, or they did not. Either oak trees and water lilies have a more recent common ancestor than oak trees have with pine trees, or they do not.

We can take Dembski's claim to be restricted to the problem of the origin of life, or we can take it to apply more generally. In this discussion I will assume that it applies more generally, not just to the origin of life but to subsequent evolutionary processes as well. So bacteria contain CSI, but camels contain more CSI than bacteria do. And I take Dembski to claim that the processes described in neo-Darwinian evolutionary theory cannot explain how we can go from bacteria to camels.

When I say "neo-Darwinian theory" I have in mind a fairly inclusive sense of this term. Neo-Darwinism combines the basic explanatory structure outlined by Darwin with modern genetics and molecular biology. Sometimes "neo-Darwinism" is also associated with strong claims about the size of mutations and the constancy of the rate of evolution, but I am assuming a flexible attitude towards these details. So I take moderate unorthodoxies such as Stephen Jay Gould's view (1989) to be neo-Darwinian in this inclusive sense.

We must explain both the origin of life and its subsequent development. I will not discuss the origin of life in detail here. There is not the kind of consensus about the origin of life that we find about the broad outlines of subsequent evolution. The whole area is more speculative, and it is tricky for a Darwinist because we are discussing the appearence of the mechanisms that make later episodes of mutation and natural selection possible. A classic experiment by Miller and Urey in 1953, and its subsequent elaboration, has shown that many of the organic molecules fundamental to living systems, including some seemingly "improbable" ones, can be created from inorganic raw materials spontaneously in laboratory experiments that simulate early

conditions on earth. The molecules produced include amino acids, which make up proteins, and also the bases that are key components of DNA and RNA. Short chains of amino acids ("proteinoids") have also been produced from solutions of individual amino acids with the aid of heat and the right kind of surface. Once we leave the laboratory results, however, most issues concerning the origin of life are very controversial. General theories divide roughly into those positing a single type of molecule with special properties of self-replication, possibly a form of RNA, and theories positing self-reproducing chemical networks, in which no single molecule is central. The appearence of cell-like encapsulation of self-replicating structures is also crucial. Discussions of the origins of life rapidly become technical. Here I will mostly discuss the role of natural causes in life's subsequent evolution. But many of the points I will make below apply also the question of the origin of life.[8]

So how can you get from a bacterium to a camel? A large part of the answer is: slowly! Neo-Darwinism explains a process of this kind by breaking it down to a large number of small steps. The exact size of the important steps, and the variation in step-size, are famously controversial issues. But there is consensus about the important *physical* features of these processes, and about the fact that *many* of these steps are needed for large-scale evolutionary transitions. Genes mutate, as a consequence of molecular mishaps. Organisms have their structure and behavior affected by the mutations, usually for ill but occasionally for good. The organisms live, reproduce and die, and those carrying novel genes either reproduce more or less than other organisms in the population. If they reproduce more, and certain other conditions are realized, the frequency of those genes in the population will tend to increase. Through this process, useful modifications slowly accumulate. Genetic material is duplicated within the genome, and the duplicates acquire new roles, making more complex structures possible. Populations change over time, split, and diverge. The striking features of evolution – elephants from bacteria, roses related (distantly) to tuna, and so on – are a consequence of the accumulation of a great many of these small steps.!I talk of "steps" here, but it is important not to have in one's mind a single linear chain of events. The local events that give rise to large-scale evolutionary processes are scattered, as the individual organisms of all the various species are. Organisms live and die locally. But a total family tree connects life on earth – connects all individuals as well as all species.

If there is more "complex specified information" in a camel than in a bacterium, then the natural process described above is able to create this information. There is

no metaphysical mystery in such a creation; consider what is required for natural causes to take us from a simple bacterium to a more slightly more complex, well-adapted and "unlikely" bacterium. What this requires is a molecular accident in the replication of the bacterial DNA that happens to result in a change to some protein molecule produced by that DNA, a change which happens to help the bacterium survive and reproduce – perhaps, for example, by enabling it to break down a environmental chemical that was formerly toxic to it. The innovation will proliferate, if circumstances allow, and in time we will have a bacterial population able to do something novel and adaptive.

A case like this is a very low-level case of Darwinian evolution, a far cry from the evolution of human eyes or brains. But the opponents of Darwinism do have to deal with these cases. Anti-Darwinians have two options here. One option is to concede these low-level cases to the Darwinist, but claim that these are not the important cases. The other option is to deny even the low-level cases.

The first option is the smarter choice for the anti-Darwinian: Darwinism, on this view, can give bacteria new adaptations to their environments but cannot bring about "irreducible complexity," and cannot generate the big transitions between whole phyla of organisms. If this option is taken, the battle must be fought over the question of how different the low-level cases are from the big and striking cases. In particular, are the largest and most dramatic evolutionary transitions explainable in terms of long chains of lower-level evolutionary events? This first option is the smarter one because it does not commit the anti-Darwinian to an absurd fight about whether bacteria can acquire antibiotic resistance or moths can get darker wings through Darwinian processes. It is also a smarter option because every time a Darwinian explains how an evolutionary process could create some particular noteworthy structure, and explains why the traces of evolutionary history are in fact visible in that structure, the anti-Darwinian can concede the case (or just stop talking about it) and pick a new one. (For this sort of debate, see the exchange between Behe, Orr, Doolittle, Futuyma, and others in the *Boston Review* issues for December/January and February/March 1997.)

Although this hedged option is the sensible one for the anti-Darwinian, it does not seem to be available to Dembski. This is because Dembski expresses his claims in the form of a "law," the "Law of the Conservation of Information." The law is expressed in sweeping terms. Natural causes cannot create *any* new CSI, not even a little bit; "the CSI in a closed system of natural causes remains constant or

decreases". So he must insist that natural causes cannot add even a tiny increment of CSI to a population of bacteria.

The exact scope of this "law" is hard to determine because, on Dembski's own view, there is no sharp cut-off between "complex" information and information that is not complex. The law does not say that information in general is conserved. But Dembski clearly thinks that something very important cannot be created, even in small amounts, by natural processes; I pass over the problem of what exactly this is.

So Dembski claims that even small additions of CSI via natural causes are impossible. He grounds this claim on an abstract argument about "chance" and "necessity." Dembski holds that neither of these, in principle, can create CSI, and a combination of the two can fare no better. But mutation and selection is, Dembski says, just a combination of chance and necessity.

Dembski misdescribes the structure of neo-Darwinian explanation in a number of ways. I will quickly mention a terminological oddity first. Dembski correctly says that the process of mutation and selection resembles learning by trial and error. But Dembski expresses the analogy by likening mutation to "error" and selection to "trial." This is a strange flipping of the usual way of describing the analogy. Usually people writing about evolution say that mutation is a trial (in the sense of a "trying-out," or an experiment) and selection corresponds to error (feedback from the environment, usually negative). In this way, a Darwinian process is seen as searching a space of possible structures. I suppose the terms "trial" and "error" always had within them the possibility of a switch from their usual roles in the analogy. But Dembski is the first person I have ever seen make this switch. Maybe he meant to do it, and a terminological oddity is certainly not a big problem. But this does illustrate how far Dembski's discussion is from the large literature discussing how natural selection actually works.

Turning to the argument itself, Dembski says we can discover some limitations of Darwinism by looking at the limitations associated with chance and with necessity. It is true that Monod (1971) and others have described Darwinism as a combination of chance and necessity. But this is a metaphorical description of Darwinism that has little value and much potential to mislead, as it does here. Mutation of genes can be treated like a random event for the purposes of much evolutionary discussion and modeling, but mutation is of course a physical process. The main reason to use the term "chance" has been to stress that mutation occurs haphazardly in many directions, and occurs in ways that are not controlled by the needs of the organism. A

14

mutation, however, is a certain type of change to a certain type of molecule within a cell. This process results in one chemical compound being substituted for another within a particular larger molecule. These molecular events are part of the ordinary causal order. Such processes might be irreducibly chancy in the sense associated with quantum mechanics – but only if everything in the world is irreducibly chancy in the same sense.

Referring to mutations as "chance" events is often a useful shorthand, but referring to selection as a kind of "necessity" has no positive use. I think people like to call selection a kind of "necessity" just because necessity is a complement to chance, and the pairings of chance/mutation and necessity/selection look neat. But selection, too, is a type of causal process (Sober 1984). This process involves the interaction of organisms with their environments, and involves in particular the various causes and effects of individual reproduction. The physical processes involved in selection have as much determinism or necessity at the lowest level as basic physics allows – just as mutation does. So the processes of natural selection are not really any more "necessary" than mutation is. Selection processes are in some respects more patterned and regular, but they are not especially "necessary." Consider our evolving bacterial population again. The mutations in bacteria all have their specific molecular causes. The mutations result in different protein molecules being produced, which affect the metabolic processes of each of the mutated bacteria. Most mutations are bad, but every now and then one turns out to be useful.

Monod's metaphorical description of evolution in terms of chance and necessity will not bear any weight, and certainly will not support an argument like Dembski's. There is no way one can deduce any important properties of evolution by working from a discussion of the general properties of chance and the general properties of necessity.[9]

More generally, when working out what evolution can bring about, it is a mistake to try to deduce strong conclusions from the basic properties of information, contingency, chance, necessity, probability and so on. All these concepts have only slippery applications to the actual processes of evolution as understood by Darwinians. They provide picturesque ways of talking, but they will not support inferences of the kind Dembski engages in. To work out what evolution can and cannot do, the way to proceed is to focus on the biological processes themselves – the variation found in populations, the phenomena of heritability and their underlying genetic mechanisms, and the interaction of organisms with their physical and

biological environments. When we look at how these processes in fact work, we find that producing outcomes that would otherwise be unlikely is *exactly* what Darwinian processes *are* able to do. The causal processes of mutation and natural selection provide a way for a population of organisms to accumulate useful bits and pieces of biological machinery – the products of very occasional good luck in mutation's products. The hard problem for Darwinism has always been showing that the biggest transitions in evolution can be explained in these terms. (For a fascinating attempt to tackle these hardest problems, see Maynard Smith and Szathmáry 1995).

In this section I have argued that his sweeping formulations and abstract claims about chance and necessity make Dembski's argument particularly vulnerable, even by creationist standards. The role played by Dembski's "law of the conservation of information," which is a vague relative of real laws in thermodynamics, exemplifes this most clearly. In these respects Dembski is, ironically, a little bit like Herbert Spencer (1871, 1872). Spencer was an English speculative polymath, influential during the late 19th century, who developed a grand theory of evolution that applied to everything, from the cosmos through to human society. Spencer had some interesting ideas, but one of his pitfalls was his tendency to express these ideas in the form of grand, general laws, which were sometimes vague analogs of Newtonian laws. Spencer had a particular fondness for conservation principles. This led him into hopelessly sweeping formulations of his claims. None of Spencer's laws are taken seriously today.

4. Conclusion

In section 2 of this paper I outlined Dembski's use of information theory. My main claim in this section was that Dembski's "complex specified information" is basically just improbability. Something contains CSI if it is very improbable for that thing to be produced via chance processes, where this improbability is not of a kind dependent on hindsight. Dembski sees living organisms as improbable things in this sense, and although there are problems surrounding this kind of application of probability theory, quite a few Darwinian biologists have expressed similar ideas. Dembski's concept of "complex specified information" does not have much to do with complexity as ordinarily understood, or as understood by biologists. Complex structures (in a biologist's sense) might indeed tend to be improbable, but plenty of improbable things are not especially complex (like winning poker hands). Dembski's use of information theory is minimal and raises no new problems for Darwinians.

In section 3 I looked at Dembski's claim that natural processes, especially those described by Darwinians, cannot generate CSI. I argued that Dembski's use of a "conservation law" to express this idea lands him in the uncomfortable position of claiming that even slight increases in complexity or adaptedness cannot be produced by Darwinian processes. I also argued that there is no way to deduce any important properties of evolution from abstract discussions of "chance and necessity." Mutation and selection are causal processes, and they are causal processes that enable populations of organisms to discover and accumulate useful new biological traits, gradually leading to large scale evolutionary change.

Dembski's paper is part of a recent movement aimed at defending a place for "intelligent design" in the overall structure of the universe. These defences include erroneous but well-packaged attacks on the explanatory power of Darwinism, and the publicity surrounding these attacks makes it harder to educate people about evolution. Understanding how evolutionary processes work is not just a matter of theoretical, academic interest. For example, in this paper I have sometimes used examples involving bacteria. There are lots of reasons to be interested in bacterial evolution, and one is the fact that the history of human disease is made up in part by the evolutionary history of disease-causing microbes (Ewald 1994, Diamond 1997). Bacteria and viruses evolve much faster than we do; they evolve on the rapid time-scale of human social change. As a consequence, we can study their evolution readily, but we are also affected by their evolution in more mundane ways. Why is the routine over-use of antibiotics a bad thing? It is bad for Darwinian reasons. The more often each antibiotic is used, the more selection pressure is put on bacteria to evolve ways to survive the onslaught from that drug. And evolving resistance is exactly what they tend to do. Over recent decades bacteria have found ways to break down, pump out, or in other ways neutralize a great range of antibiotic drugs. Some have argued that bacteria are *so* good at this sort of thing that Darwinian explanations of bacterial evolution must be supplemented with "directed mutation." (See Jablonka and Lamb 1995 chapter 3 for a survey of these issues). The jury is still out on the issue of whether there is any role for a significantly non-Darwinian form of evolution in bacteria, but even those defending unorthodox views in this area do not doubt that a great deal of bacterial evolution is straightforwardly Darwinian.

Hospitals are where many of these dangerous bacteria are commonly found, both because hospitals are full of sick people and because hospitals are special environments where the only bacteria that can survive are the ones which can beat

17

the chemicals we use against them. When a mutant bacterium appears in a hospital which happens to be able to survive in the presence of available antibiotics, the new strain will often multiply and we will be the worse for it (Garrett 1995). There is no way to avoid altogether this problem with antibiotics; given the nature of bacteria and the nature of evolutionary processes, we must expect resistance to arise eventually in many cases. But knowledge of microbial evolution instructs us to use antibiotics only when there is good reason to do so. The drugs will inevitably be more effective the less they are used. These are Darwinian facts, and facts with immediate practical importance.

In sum: the classical argument from design was answered by Darwin and by the subsequent development of evolutionary biology. Recasting the argument in terms of "information" does not change that situation. And a recasting in terms of a general "law of conservation of information" makes the argument worse rather than better.

*   *   *   *   *

Notes
[1] Dembski's paper "Intelligent Design..." is my main target in this discussion. Dembski has also published a book (1998) which develops his claims about probability in more detail, and which includes a brief discussion of the evolution/creation controversy (pp. 55-62). The discussion of evolution and creation in the book does not take sides. In footnotes to this paper I will indicate a few areas in which Dembski's book and his article differ on points of relevant detail. None of these issues are important to the overall argument in "Intelligent Design...," and my criticisms of that argument.

[2] Thanks to Amir Najmi for assistance with these issues. Shannon (1948) did refer to the source as a "source of information," although he was considering the source in relation to a signal and a receiver. Through the rest of this paper I will use the term "information" in the way Dembski does, without further comment on the relation between the terms "information," "entropy" and so on.

[3] In this paper I use the term "correlation" in a broad vernacular sense, not in the strict sense measured by a correlation coefficient. I also use the terms "probable" and "likely" interchangeably; "likely" does not refer specifically to likelihoods, in the statistician's sense.

[4] In his 1998 book, Dembski focuses more on probabilities and does not dress his argument up in information-theoretic terms.

[5] In his 1998 book, Dembski discusses complexity theory in more detail. I am not qualified to comment on technical issues in complexity theory. But, again, the only role played by "complexity" in his argument in "Intelligent Design..." is as a measure of low probability.

[6] Dembski quotes Dawkins and notes one aspect of their agreement, though he expresses Dawkins' claim as if Dawkins was making a concession to Dembski's position.

[7] Dembski expands on his positive views, though not on biological matters, in a paper called "The Act of Creation: Bridging Transcendence and Immanence," which is available at http://www.arn.org/docs/dembski/wd_actofcreation.htm.

[8] See Maynard Smith and Szathmáry (1995) for a critical survey of recent work on the origin of life. A standard college biology textbook, Campbell, Reece and Mitchell (1999), also has a very good overview, pp. 492-498.

[9] Dembski's book makes use of a classification in which there are only three possible kinds of explanation for any event – regularity, chance and design (1998 chapter 2). But many events, and a great many of the ones relevant to these debates about evolution, result from a complicated causal network which includes both likely and unlikely events, interacting over time to yield an outcome. A simple three-way distinction between regularity, chance and design is entirely inadequate for understanding explanation in such cases.

References

Bonner, J. T. (1988). *The Evolution of Complexity*. Princeton. Princeton University Press.

Campbell, N. A., J. B. Reece, and L. G. Mitchell (1999). *Biology*, 5th edition. Menlo Park: Addison Wesley.

Cover, T. M. and J. A. Thomas (1991). *Elements of Information Theory*. New York: Wiley.

Dawkins, R. (1986). *The Blind Watchmaker*. New York: Norton.

Dembski, W. (1997) "Intelligent Design as a Theory of Information" *Perspectives on Science and Christian Faith* 49 (3): 180-190, reprinted in R. Pennock (ed.) *Intelligent Design Creationism and its Critics: Philosophical, Theological and Scientific Perspectives*, MIT Press, 2001.

Dembski, W. (1998) *The Design Inference*. Cambridge: Cambridge University Press.

Diamond, J. (1997). *Guns, Germs and Steel: The Fates of Human Societies*. New York: Norton.

Dretske, F. (1981). *Knowledge and the Flow of Information*. Cambridge MA: MIT Press.

Ewald, P. A. (1999). *The Evolution of Infectious Disease*. Oxford: Oxford University Press.

Garrett, L. (1995). *The Coming Plague*. New York: Penguin.

Gilbert, S. F., J. M. Opitz and R. A. Raff (1996). "Resynthesizing Evolutionary and Developmental Biology." *Developmental Biology* 173: 357-72.

Godfrey-Smith, P. (1996). *Complexity and the Function of Mind in Nature*. Cambridge: Cambridge University Press.

Gould, S. J. (1989). *Wonderful Life*. New York: Norton.

Jablonka, E. and M. Lamb (1995). *Epigenetic Inheritance and Evolution*. Oxford: Oxford University Press.

Johnson. P. (1996). "Is Genetic Information Irreducible?" *Biology and Philosophy* 11: 535-538.

Maynard Smith, J, and E. Szathmáry (1995). *The Major Transitions in Evolution*. Oxford: Oxford University Press.

McShea, D. (1991). "Complexity and Evolution: What Everybody Knows," *Biology and Philosophy* 6: 303-24.

Monod, J. (1971). *Chance and Necessity*. Translated by Austryn Wainhouse. New York: Alfred A. Knopf.

Shannon, C. E. (1948). "A Mathematical Theory of Communication," *Bell System Technical Journal* 27: 379-423. 623-656.

Sober, E. (1984). *The Nature of Selection*. Cambridge MA: MIT Press.

Spencer, H. (1871). *Principles of Psychology*. 2nd edition. 2 volumes. New York: Appleton.

Spencer, H. (1872). *First Principles of a New System of Philosophy*. 2nd edition. New York: Appleton.

Williams, G. C. (1992). *Natural Selection: Domains, Levels and Challenges*. Oxford: Oxford University Press.