Three Issues With "No Free Lunch"

Darel R. Finley

In *No Free Lunch,* (Rowman & Littlefield 2002, ISBN 0-7425-1297-5) William Dembski lays out his thesis of CSI (complex specified information), and identifies it as a reliable indicator of intelligent action. Some of the book is a re-presentation of previous material, but in a more widely readable form, although even in *No Free Lunch* there are portions of the book that become eye-glazingly technical, and perhaps would have benefited from more analogies and/or examples. Symbology-laden sections in chapter 2 can be skimmed without losing the main message of the book, although I warn the reader not to skim section 3.1.

No Free Lunch covers many important topics, including:

* Complex Specified Information (CSI), and information as improbability
* the probabilistic resources available in our universe, and the fallacy of speculatively inflating them
* the explanatory filter, and criticisms of it
* the law of conservation of information, as a proposed Fourth Law of Thermodynamics
* evolutionary algorithms and the displacement problem
* Behe's concept of irreducible complexity refined, and defended from counterarguments
* Intelligent Design as an argument from *invariance,* not from ignorance
* questions posed by Design as a research program

Dembski's book is undoubtedly an important addition to the growing set of books in the Intelligent Design movement, and I wholeheartedly recommend it to anyone interested in that movement. However, when I finished reading the book I was left with the uneasy feeling that a few important points were being overlooked; hence, this paper. I present my points in a non-technical form, based primarily on an analogy to finding a specified square on a large grid. However, I have little doubt that these issues are valid, and that applying them to biology is ultimately unavoidable.

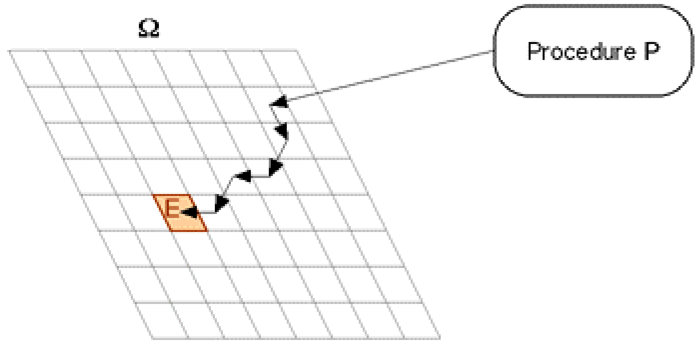
Issue 1 — *Dual Definition of Probability*

After spending much of chapter 3 explaining that information is equivalent to improbability within a phase space of all that might have been (such as an arrow hitting a particular spot on a wall, as opposed to all the other spots it might have hit instead), Dembski then says on page 183 (chapter 4):

> But a probability amplifier is also a *complexity diminisher.* For something to be complex, there must be many live possibilities that could take its place. Increasingly numerous live possibilities correspond to increasing improbability of any one of those possibilities. Complexity and probability therefore vary inversely — the greater the complexity, the smaller the probability. It follows that Dawkins's ["me-thinks-it-is-like-a-weasel"] evolutionary algorithm, by vastly increasing the probability of getting the target sequence, vastly decreases the complexity inherent in that sequence. As the sole possibility that Dawkins's evolutionary algorithm can attain, the target sequence in fact has minimal complexity (i.e. the probability is 1 and the complexity, as measured by the usual information measure, is 0). Evolutionary algorithms are therefore incapable of generating true complexity. And since they cannot generate true complexity, they cannot generate true specified complexity either.
>
> In general, then, evolutionary algorithms generate not true specified complexity but at best the *appearance of specified complexity*. This claim is reminiscent of one made by Richard Dawkins. On the opening page of *The Blind Watchmaker* he states, "Biology is the study of complicated things that give the appearance of having been designed for a purpose." Just as the Darwinian mechanism does not generate actual design but only its appearance, so too the Darwinian mechanism does not generate actual specified complexity but only its appearance.
>
> [Emphasis in original.]

This passage seems to be trying to define the probability of an event two different ways simultaneously — as both the fraction of the phase space covered by the event, and as the likelihood that a given procedure will produce the event. For example:

Figure 1: Defining Probability



Ω = phase space (squares on a chessboard)

E = a specified event

P = an iterative procedure that is 100% certain to find **E** in relatively few iterations

Definition A:

"probability of **E**" $=_{def}$ degree of specificity of **E** within Ω
= 1/64
= 6 bits of information

Definition B:

"probability of **E**" $=_{def}$ certainty with which procedure **P** will arrive at **E**
= 1
= 0 bits of information

Definitions A and B arrive at different measures of information for **E** (6 vs. 0 bits), and are thus mutually contradictory and cannot be used interchangeably. Dembski, however, seems to be using them interchangeably when he asserts that any event **E** arrived at via an evolutionary algorithm can contain only the *appearance* of information; i.e. what appears to be 6 bits is really 0 bits. If we apply that logic to the neo-Darwinian synthesis, we might conclude that the human genome only appears to contain huge amounts of specified complexity, but really contains 0 bits (if evolution was certain to create it) or a modest number of bits (if evolution might have created $2^{\text{modest number}}$ of variations on the human body, and happened by chance to produce this one). But that would scarcely matter to evolutionists — whether you call it 0 bits or not, if random mutation and natural selection produced the human genome, then evolution is correct.

Given what follows on the next several pages, I suspect that what Dembski was really trying to say on page 183 (as applied to my "Procedure P" example) is that event **E** really does contain 6 bits, but that procedure **P** produced 0 of those bits, and only appeared to produce 6. That appearance was created by the fact that the 6 bits of specified information were front-loaded into the procedure **P**, or into a fitness function **j** that **P** exploits. And in any case, the dual definition of probability seems to be confined to page 183 and is not a crucial linchpin of the book's most important arguments.

Issue 2 — *Persistence of the Displacement Problem In A Non-Evolutionary Biology*

To explain this issue, let me begin with an example of the displacement problem in action. Suppose we have a very large chessboard, 2^{250} by 2^{250} squares on each side. Thus, the whole board has 2^{500} squares (more than 10^{150}). This means that all the probabilistic resources our universe has to offer cannot render reasonably probable the selection of a particular (independently specified) square at random, as detailed in section 1.5 of *No Free Lunch.*

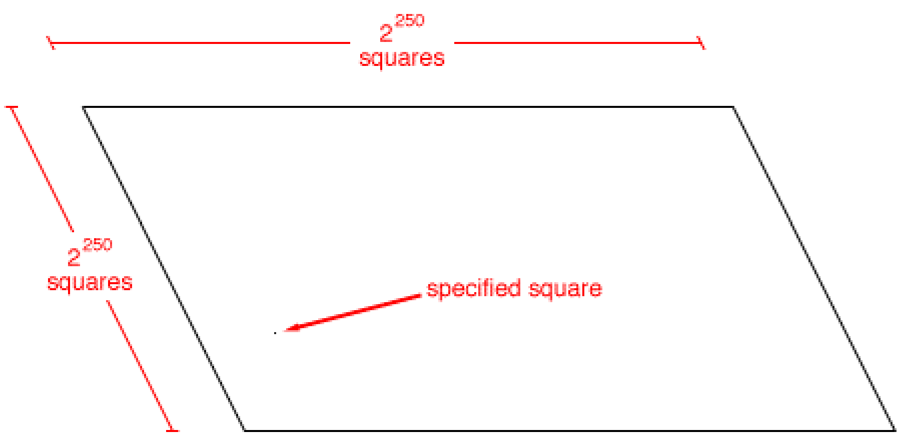


Figure 2. A Large Phase Space

Now suppose that one of these squares is interesting to us, for some reason independent of whether the square is currently selected/instantiated. For example, the square pointed to by the arrow in Figure 2 is such that if you take its Y coordinate (a 250-bit integer), string it together with its X coordinate (another 250-bit integer), and then consider the whole 500-bit sequence to represent a 100-character sequence, where each group of 5 bits corresponds to a single character from the following set of 32 characters,

“**ABCDEFGHIJKLMNOPQRSTUVWXYZ .,!?-**”,

where 00000 = A, 00001 = B, 00010 = C, etc., then the whole 100-character sequence happens to spell out this phrase:

“**WASHINGTON STATE IS SEPARATED INTO TWO DISTINCT AND DISSIMILAR ZONES BY A SPINE OF RUGGED MOUNTAINS.**”

Figure 3 shows the message decoding method in action:

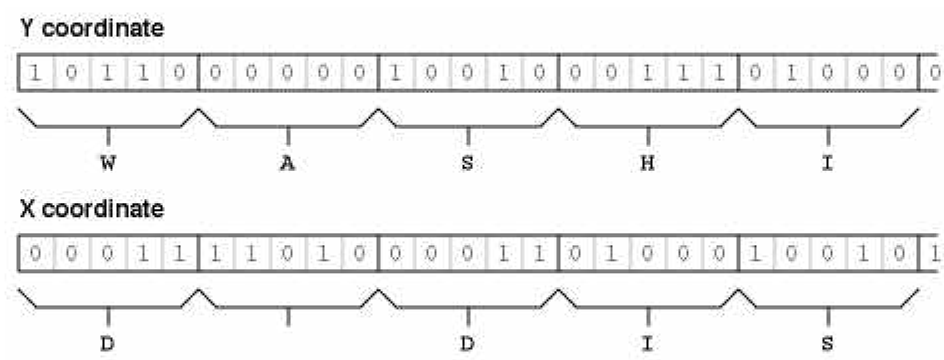


Figure 3. Five-Bit Character Encoding Scheme

Now let’s suppose that we want to write a computer algorithm that will find that particular square (and thus its associated phrase about Washington geography). How could we do it? The most direct way would be:

The GO-TO Algorithm
1. Select the desired square immediately.
2. Stay there.

Figure 4. The GO-TO Algorithm

The GO-TO algorithm is unsatisfying, because it requires that the 500 bits of data be preloaded into the algorithm, so there is no point in even running the algorithm. What we want is an algorithm that will find the specified square *without* having to be pre-loaded with the square's coordinates. Here are two algorithms that do exactly that:

The RASTER Algorithm
1. Select the square in the upper-left corner (0,0).
2. Advance one square to the right.
3. Repeat step 2 until reaching the end of the row.
4. Move back to the beginning of the row, and down one row.
5. Go back to step 2.

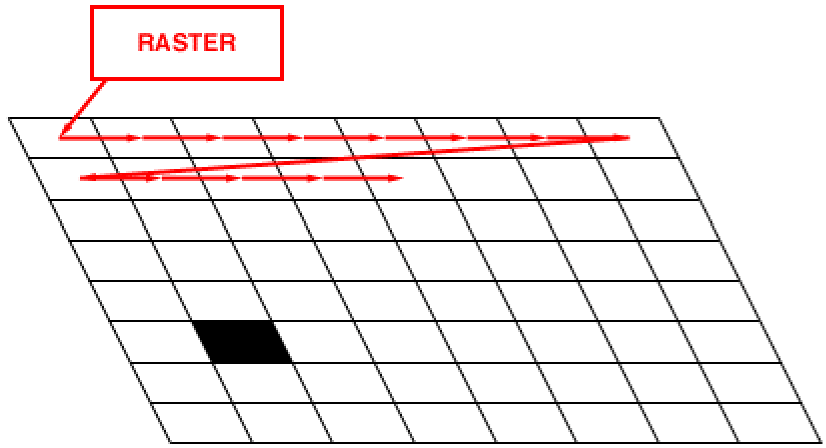


Figure 5. The RASTER Algorithm *Simplified to 8x8 for illustration purposes.*

The RASTER algorithm scans the chessboard in reading order, much as the scanning beam in a CRT monitor paints the picture onto its phosphor screen.

The RANDOM Algorithm
1. Select a square at random.
2. Go back to step 1.

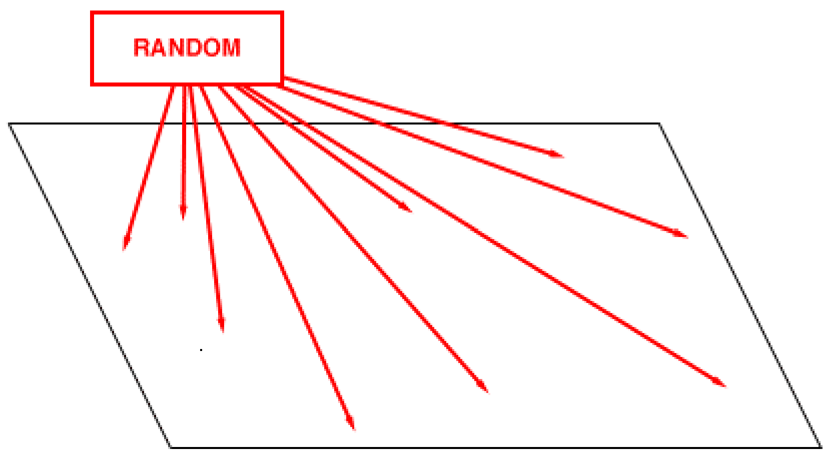


Figure 6. The RANDOM Algorithm

The RANDOM algorithm is guaranteed to eventually hit our specified square by chance.

The RASTER and RANDOM algorithms do not contain the 500 bits of data pre-loaded (as did the GO-TO algorithm), and they will arrive at the specified square, but they suffer from the fatal problem of taking much too long to execute. Even on a very fast computer, our universe will not provide enough time for the RASTER algorithm to reach the specified square, and not enough time for RANDOM to have more than an infinitesimally small chance of hitting it. Thus, neither algorithm has any practical value for finding specified data like the "Washington state" phrase.

How about an evolutionary algorithm? Suppose we have a fitness function J that is shaped like a simple curve that rises to a peak, and it happens to peak at our specified square. It would look something like this:

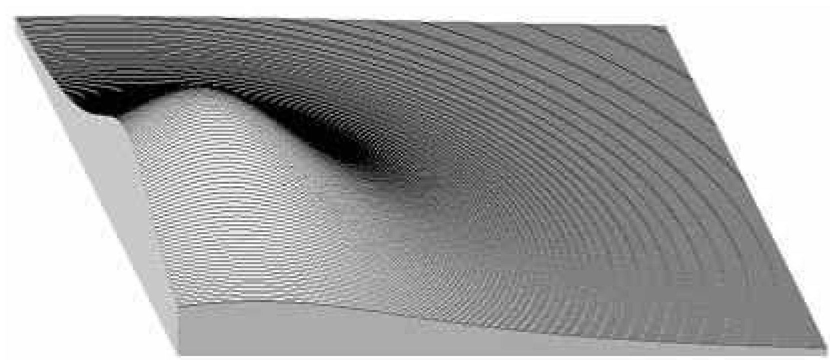


Figure 7. A Fitness Function

Now we can write an evolutionary algorithm that, guided by fitness function J, will find the data in a reasonable time:

The EVOLUTION Algorithm
1. Select a square at random.
2. Test a randomly selected nearby square to see if it scores higher on function J.
3. If the test square scores higher, move to it — otherwise stay put.
4. Go back to step 2.

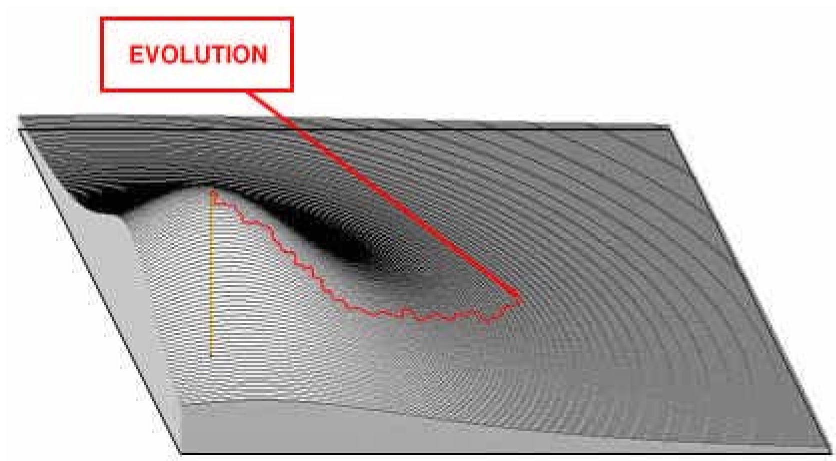


Figure 8. The EVOLUTION Algorithm

If this algorithm is written properly, it will, by a meandering path, reach our specified square in a very reasonable amount of time, while we are still around to see it happen. And in doing so, it seems to convincingly demonstrate that evolution is a good theory and probably does work. What might be wrong with it?

Fitness Function Roughness

A common argument that is leveled against evolution is that the real-life, biological fitness function (over DNA sequence space) doesn't slope up nicely as pictured in Figure 3, but instead has sharp buttes of viability, separated by wide expanses of sterile flatness.

Figure 9. A Rough Fitness Function For the "Washington" Square

The flat regions are not only unable to guide an evolutionary process to any meaningful data, but additionally are *lethal* — they immediately terminate any evolutionary process that wanders into them.

However, the CSI-displacement problem is not concerned with the practical matter of a rough fitness function — so let's assume, for the moment, that the function is a nice smooth curve as pictured in Figure 7. In that case, is there any *other* reason to doubt that the EVOLUTION algorithm can generate our "Washington" phrase?

CSI Displaced To Fitness Function

There is. What sort of function generates a nice, smooth curve that rises to a peak? It doesn't take a very complex formula to do it:

Function 1: $\mathbf{J = 1 / (1 + x^2 + y^2)}$

Surely a simple function like that can be expressed in less than 500 bits, and since we can make our chessboard as large as we like (requiring coordinate data of a thousand bits, or a million bits, etc.), our evolutionary algorithm is clearly generating large amounts of specified data from a very simple formula.

But Function 1 actually peaks in the *upper-left corner of the chessboard,* which represents the 500 bits being all set to *zero.* Graphed, it looks like this:

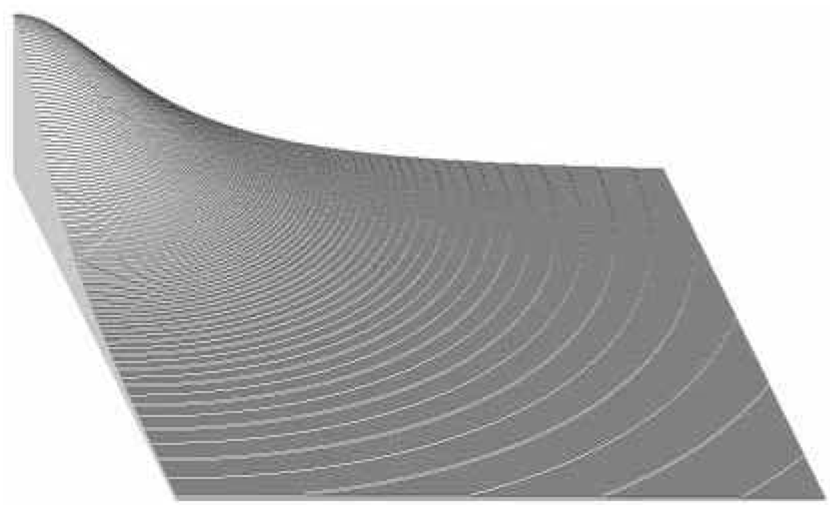


Figure 10. Function 1 Peaking At Useless Data

Function 1 is indeed simple, and it generates very simple data — all zeros. It does not generate CSI. 500 bits of all zeros, translated through the encoding scheme in Figure 3, reads:

> "**AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA**
> **AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA**"

And no simple encoding scheme will translate it to anything complex and useful. To make the function peak at the "Washington" square (as in Figure 7), we will need to modify it as follows:

Function 2: $\mathbf{J = 1 \,/\, (1 + (x\text{-}M)^2 + (y\text{-}N)^2)}$

where M and N are the coordinates of the specified square. For J to peak at the correct square, M and N must each be specified to an accuracy of 250 bits — therefore, Function 2 has the *whole 500 bits of specified information pre-loaded into it,* before the EVOLUTION algorithm even starts running! That requirement makes EVOLUTION no better than

GO-TO — both require the whole phrase to be pre-loaded. (This is functionally identical to Richard Dawkins's "methinks it is like a weasel" program, which has to have the whole target phrase pre-programmed into its fitness function to make the algorithm work.) So the question "Where did the CSI come from?" is not answered by invoking an evolutionary algorithm — that invocation merely *displaces* the CSI question to the algorithm's fitness function; i.e. "Where did the CSI-loaded fitness function come from?"

The formal way to describe it — keeping with the framework laid out by Dembski — is that to contingently select Function 2 from the larger phase space of all possible fitness functions over our super-chessboard, is to select the full CSI of the "Washington" phrase, in advance of any evolutionary algorithm finding it with the aid of the function. A more practical way to describe it is simply to note that any computer implementation of Function 2 would have to include the full 500 bits of the "Washington" phrase in its compiled (launchable) code, whereas a computer implementation of Function 1 would not.

<u>*Holding On To CSI*</u>

Now lets look back at the RASTER and RANDOM algorithms. They take much too long to be useful, but at least they find the specified square without having its coordinates pre-loaded, right? Well, not exactly. Yes, RASTER and RANDOM will eventually hit the specified square, *but then they will leave it immediately* and move on to another square. Even if your universe happened to have unlimited amounts of time to wait, what good is an algorithm that hits interesting data for a tiny fraction of a second and then immediately loses it? Valuable genetic information needs not only to be found, but also kept and used. Only GO-TO and EVOLUTION will hold on to the data, once found, and they are both pre-loaded with the entire target phrase.

If evolutionary algorithms depend on fitness functions that must be pre-loaded with the data that the algorithm is going to "generate," then the standard evolutionist claim — that evolution designs complex systems and synthesizes large quantities of useful genetic data — is false. For biology, the implication of the displacement problem is that either

1. The laws of physics (which determine the fitness function) are pre-loaded with all the useful, complex data (CSI) that is currently found in the DNA of living organisms, and provide an evolutionary path to realizing that data in actual organisms, or

2. The laws of physics contain a relatively small amount of data — nowhere near the amount needed to make a human — and so the human DNA was written by a designer, who inserted large amounts of specified information (CSI) into the universe at various points in the history of life on Earth.

Dembski shows in Chapter 5 of *No Free Lunch* that scenario 1 cannot be true, due to the irreducible complexity of many biological structures; a phenomenon also described as the "holistic" nature of the information that specifies such structures. Thus, scenario 2 is inferred.

But notice that scenario 1 is actually a compound assertion; i.e. "laws are pre-loaded **and** provide an evolutionary path." A compound assertion, in a list of two purportedly comprehensive options, suggests the possibility of a false dichotomy. What if there is a third option, as follows:

3. The laws of physics are pre-loaded with all the CSI that is currently found in the DNA of living organisms, but do not provide an evolutionary path to realizing that data in actual organisms.

In practice, scenario 3 would go like so:

3a. The designer starts a universe, giving it a set of physical laws that are pre-loaded with all the CSI that organisms need, but offers no evolutionary path to natural formation of the organisms (perhaps because a smooth fitness function is not feasible in an interesting universe).

3b. After letting the universe develop naturally to a certain point, the designer begins creating organisms on a suitable planet by direct intervention.

3c. The organisms survive and flourish into large populations.

3d. Over time, the organisms' DNA code is damaged by mutation, but the naturally occurring fitness function (determined by the physical laws) allows natural selection to weed out the errors and keep the species on its local fitness butte.

Given that scenario 1 is not correct, what reason is there to think that scenario 3 is more likely than scenario 2? Look again at Figure 9, the rough fitness function for the "Washington" CSI:

Figure 9 (duplicate). A Rough Fitness Function For the "Washington" Square

What sort of fitness function J would make a sharp peak like that? Here's one:

Function 3: **If (x = M and y = N) then J = 1, else J = 0**

And here's another:

Function 4: $\mathbf{J = 1 / (1 + K(x\text{-}M)^2 + K(y\text{-}N)^2)}$

where K is a very large constant, say 10^{200}. Functions 3 and 4 will generate a fitness function with a sharp spike at M,N, but they both have the full 500 bits of data (M,N) pre-stuffed into them, just as did Function 2. This is consistent with our earlier observation that only the pre-stuffed algorithms could *hold on* to CSI.

If my thesis (Scenario 3 above) is correct, then one or more of the following must be true:

a. The laws of physics are more complicated than popularly perceived, and contain the full complexity of organismal CSI. (The laws include not just the laws of relativity and quantum mechanics, but also everything required to make atoms behave and interact in the ways they do.)

b. The amount of CSI in living organisms is substantially less than it appears, perhaps being confined to a very small subset of an organism's genome.

c. The CSI of living organisms may be fragmentable into small, independent pieces of CSI, which could be combined in many different ways to make functional organisms.

d. Some of the CSI in organisms may be specified only as it interacts with other CSI, and thus those two (or more) items of CSI, taken together, do not constitute true CSI (as related to the laws of physics). As an analogy, consider the particular shape of two interlocking pieces of a jigsaw puzzle. Each is highly specified in that it fits the other, but taken together there is no real specification; only a randomly curved interface.)

Obviously, the issue is complex and there are many possibilities to be worked out as more is learned about the content and functionality of the genomes. But one thing seems clear: If the fitness-function-CSI-displacement issue is a problem for the evolutionary scenario, it cannot be waved off for the non-evolutionary scenario, because whether evolution is true or not, all of the following are unquestionably real:

* There is a DNA phase space.

* There is a real-life fitness function over DNA phase space that determines which sequences of DNA make functional organisms and which don't. Some parts of the fitness function fluctuate with the environment, but other parts do not, since they deal with tasks that don't change over time, like extracting oxygen from the air, or propelling an organism through water.

* The static portions of the real-life fitness function are directly determined by the laws of physics.

* Mutations create random changes to the DNA of all living organisms.

* Natural selection, employing the real-life fitness function, preserves the CSI in the DNA.

In light of the above facts, it should be apparent that the issue of CSI displacement to a fitness function is not just a device for refuting evolution, but is a concept that must be reconciled with any theory of biology, evolutionary or non-evolutionary.

Issue 3 — *CSI As A Possible Source of CSI*

Dembski identifies intelligence as the source of CSI, but what exactly is intelligence? Dogs are undoubtedly not *as* intelligent as humans, but are they nevertheless an intelligence? Do dogs leave effects that cannot be explained except as the result of intelligence? These sorts of questions are not explored in *No Free Lunch.*

What if intelligence is just an information processing engine (i.e. a Turing machine) with some amount of CSI pre-loaded into its memory? In that case, even a very simple device like a cell-phone is intelligent, though much less so than a human. With intelligence defined this way, CSI should be considered merely a reliable indicator of *previously existing CSI,* and a mechanism for transmitting or translating it.

And, with intelligence defined this way, consciousness would be a separate phenomenon from intelligence. So an entity could be intelligent and conscious (e.g. a human), intelligent and unconscious (e.g. a personal computer), unintelligent and conscious (i.e. a hypothetical entity that consciously experiences joy, pain, or other sensations, but has no memory or capacity for logical analysis), or unintelligent and unconscious (e.g. a rock).

In conclusion: CSI, the displacement problem, and the law of conservation of information are very important topics to the subject of origins, but their full implications have yet to be worked out, and *No Free Lunch* leaves important areas undiscussed. To Dembski's list of research problems (pp. 312-313), I would add the following; otherwise the book is a fine read and a welcome step toward incorporating information theory into the Intelligent Design school of thought.

* Which parts of the existing genomes are pre-specified by the laws of physics (for a species to be viable at all)?

* Which parts of the existing genomes are specified only in conjunction with other DNA?

* Which parts of existing sequences of (non-junk) DNA are not specified at all, and represent whim/randomness?

* Can the DNA that directs human brain wiring be shown to contain the full intellectual capacity of the human mind?