ON DEMBSKI'S LAW OF CONSERVATION OF INFORMATION

ERIK

ABSTRACT. Dembski's "Law of Conservation of Information" is investigated using a more formal approach than is provided in his latest book "No Free Lunch". It is concluded that the Law of Conservation of Information is mathematically unsubstantiated.

1. INTRODUCTION

William Dembski is currently one of the front figures of the so-called Intelligent Design (ID) movement. This movement sets out to replace or modify the current theories of abiogenesis and biological evolution to explicitly include the intervention of an intelligent designer. One of the characteristics of the movement is a lack of interest in providing descriptive explanations and models. Instead, the ID movement sets out to establish the involvement of an intelligent designer by disproving the currently available alternatives that do not require one. Dembski's ideas constitute the most general and sophisticated proposal to disprove the currently available alternatives to date. Loosely speaking, the idea is that if an event is both very unlikely with respect to a particular hypothesis and fits a pattern that in a certain unformalized sense is "independent" of the event, then we should reject that particular hypothesis in favour of a general hypothesis that simply states that an unspecified designer was involved in the realization of the event. The hypothesis of the unspecified intelligent designer is considered established if all currently available alternatives can be rejected in this fashion. In his latest book, Dembski describes his ideas in more detail, provides a subjective quasi-mathematical "formalization", and applies this "formalization" to a biological feature [1]. Thorough refutations of Dembski's entire project have already been written [3, 8]. Here I will therefore focus on a specific part of Dembski's total work; namely, his argument to establish the so-called "Law of Conservation of Information". My treatment will be as formal as possible. Dembski has been criticized for merely cloaking the simple idea that certain biological features are unlikely to be produced by anything except an intelligent designer in difficult-to-understand jargon. For instance, one reviewer of one of Dembski's previous books remarked

> "Just as there are chess players who never resign a lost game but play on until they are checkmated, one can be oblivious to one's defeat in debate, if one is clever enough to invent argument after argument, however sophistical, scoring points on minor and reparable issues, and triumphally shooting down bad arguments that were never central to the debate anyway." [4]

Date: July 6, 2002.

Perhaps by treating Dembski's ideas as formally as possible, I unwittingly encourage the unnecessary jargon and peripheral arguments that Dembski has been criticized for. Nonetheless, I feel that a formal investigation of the ideas would be valuable. Formalization is a powerful tool to detect, and perhaps also repair, logical flaws. I stress that I will not be concerned with the way Dembski's ideas are being applied to real-world cases. I, like many other critics, consider Dembski's applications to the real world to fail to conform to even the rules he set up himself. But here I will be concerned with Dembski's ideas in theory, not in practice.

Finally, it should also be mentioned that Dembski's writings are notoriously difficult to interpret. To my knowledge, in the cases where Dembski has responded to criticism, the response has always been that the critic has misunderstood him in some fatal way. My own interpretations probably differ a little from those of all other critics, but I cannot guarantee that Dembski will agree with my presentation of his ideas. The reader is encouraged to critically compare my interpretation with Dembski's original writings.

2. Notation and Definitions

Dembski's work revolves around a number of concepts original to him. The main concepts are compactly described by him as follows:

> "Given a reference class of possibilities Ω, a chance hypothesis $\mathbf{H}$, a probability measure induced by $\mathbf{H}$ and defined on Ω (i.e., $\mathbf{P}(\cdot|\mathbf{H})$), and an event/sample E from Ω; a rejection function f is *detachable* from E if and only if a subject possesses background knowledge $\mathbf{K}$ that is conditionally independent of E (i.e., $\mathbf{P}(E|\mathbf{H}\&\mathbf{K}) = \mathbf{P}(E|\mathbf{H})$) and such that $\mathbf{K}$ explicitly and univocally identifies the function f. Any rejection region R of the form $T^\gamma = \{\omega \in \Omega | f(\omega) \geq \gamma\}$ or $T_\delta = \{\omega \in \Omega | f(\omega) \leq \delta\}$ is then said to be *detachable* from E as well. Furthermore, R is then called a *specification* of E, and E is said to be *specified*." [1, pp. 62–63]

In the interest of clarity and completeness I will write out, one by one, all the relevant definitions of these concepts below. Some of them are the same as those in Dembski's book and some may differ slightly. Throughout this essay Ω will denote a sample space under consideration.

Definition 1 (Rejection region). *An event of the form* $T_{f,\delta} = \{\omega \in \Omega | f(\omega) \leq \delta\}$ *or* $T_f^\gamma = \{\omega \in \Omega | f(\omega) \geq \gamma\}$*, where* f *is a mapping* $f : \Omega \to \mathbb{R}$ *and* $\delta, \gamma \in \mathbb{R}$ *are constants, is called a* rejection region*. The function* f *is called a* rejection function*.*

In this article there will be no need to consider both forms of the rejection region, so all rejection regions will be assumed to be of the form $T_{f,\delta}$ (a very small restriction since $T_{f,\delta} = T_{-f}^{-\delta}$). It will also be clear from the context which rejection function is used, so the indices will be left out. For a given event, Dembski hints that one should choose

$$\delta = \min\{\delta' \in \mathbb{R} | E \subseteq T_{\delta'}\}, \tag{1}$$

where E is the event under study [1, p. 72]. That is, a value that makes the rejection region as small as possible should be chosen.

The arguments in [1] depend heavily on the concept of background knowledge and its relations to the sample space. Unfortunately it is left unformalized, so here I will introduce some definitions that to my knowledge do not appear in Dembski's work. In particular, rejection functions that are "explicitly and univocally" identified by background knowledge play a special role. I will denote the set of all background knowledge by $\mathcal{K}$. We also need a set of hypotheses $\mathcal{H}$.

Assumption 1. *The set $\Omega \times \mathcal{K}$ is mathematically meaningful and finite, and every hypothesis $h \in \mathcal{H}$ determines a probability measure P_h on it.*

The assumption of finiteness is just made to avoid potential technical difficulties. When more than one sample space is considered, the corresponding assumptions are of course made for all the sample spaces under study. The assumption that we can evaluate the probability of knowledge may or may not be philosophically problematic, but I will simply note that Dembski's arguments and ideas depend on it and assume that it is true for the sake of the argument. For convenience, I will write $P_h(E, K)$ instead of $P_h(E \times K)$, where $E \subseteq \Omega$ and $K \subseteq \mathcal{K}$. Also, if $E = \Omega$, I shall simply omit it and write $P_h(K)$. The corresponding convention will be used if $K = \mathcal{K}$. I will also allow myself to sometimes use an outcome, rather than the elementary event containing the outcome, as argument in the probability measure.

Dembski also requires the user of his method to determine whether background knowledge "explicitly and univocally" identifies given rejection functions. For this purpose I define a relation as follows

$$\mathcal{K}\mathrm{ID}_\Omega(f, K) = \begin{cases} \text{true} & \text{if } K \text{ explicitly and univocally identifies } f, \\ \text{false} & \text{otherwise,} \end{cases} \tag{2}$$

where f is a mapping $f : \Omega \to \mathbb{R}$ and $K \subseteq \mathcal{K}$. From a philosophical point of view one could have considerable doubts as to whether the nature of knowledge is such that it is possible to determine whether it "explicitly and univocally" identifies a function, but again I will simply note that Dembski's arguments seem to require it and assume that a relation $\mathcal{K}\mathrm{ID}$ is given in a mathematical form.

Assumption 2. *The relation $\mathcal{K}\mathrm{ID}_\Omega$ is meaningful and available in mathematical form.*

This assumption is made for all sample spaces if more than one sample space is studied.

With the notation and assumptions introduced we are now ready to state the definition of detachability and specifications.[1]

Definition 2 (h-detachability)**.** *Let $h \in \mathcal{H}$ be an hypothesis. A rejection function f is said to be h-*detachable *from an event $E \subseteq \Omega$ if there exists a $K \subseteq \mathcal{K}$ such that $\mathcal{K}\mathrm{ID}_\Omega(f, K)$ and $P_h(E) = P_h(E|K)$.*

Definition 3 (h-specification)**.** *A rejection region $T \subseteq \Omega$ defined by a rejection function that is h-detachable from the event $E \subseteq T$ is said to be a h-*specification *of E and E is said to be h-*specified.

[1]Curiously, the key concepts are defined in terms of *events* rather than *outcomes*, e.g. events—not outcomes—are specified. Perhaps this has the advantage of allowing incomplete knowledge, but it also adds considerable subtleties to the argument. For instance, if the outcome ω is observed, then all events containing this outcome must be said to have occurred. Thus, when investigating whether the outcome is such that the ID hypothesis of intelligent intervention should be inferred, one is free to consider any event that contains this outcome. If a rejection function is not h-detachable from the event $\{\omega\}$, perhaps it will be h-detachable from, say, $\{\omega, \omega'\}$.

Definition 4 (h-specified information)**.** *An ordered pair of events (E, T) such that $E \subseteq T$, and where T is a rejection region with a rejection function that is h-detachable from E, is called* h-specified information.

Roughly speaking it is h-specified information for which the h-specification is sufficiently improbable that Dembski considers to be an infallible indication that the hypothesis h should be rejected in favour of a hypothesis that an intelligent agent was involved in the realization of the event.

The rest of this essay will investigate to what extent the mathematical relations defined above are sufficient to establish Dembski's Law of Conservation of Information.

3. The Law of Conservation of Information: The Deterministic Case

Here the Law of Conservation of Information as it is described on pp. 152–153 of [1] will be investigated in detail.

Consider two sample spaces Ω and Λ and an arbitrary mapping $\phi : \Omega \to \Lambda$ between them. For convenience, define two associated mappings on the corresponding event spaces by

$$\bar{\phi}(A) = \{\phi(\omega) | \omega \in A\}, \tag{3}$$

$$\bar{\phi}^{-1}(B) = \{\omega \in \Omega | \phi(\omega) \in B\}, \tag{4}$$

where $A \subseteq \Omega$ and $B \subseteq \Lambda$.

If the mapping ϕ can be thought of as a representation of a deterministic process, Dembski's idea that a deterministic process cannot generate specified information may be expressed in terms of ϕ. There are two ways in which specified information can potentially be generated. Either an event $A \subseteq \Omega$, which is not specified, can be mapped onto an event $B \subseteq \Lambda$, which *is* specified. Or a specified event A could potentially be mapped onto an event B, such that the specification of B has lower probability than the specification of A. It must also be made clear relative to which hypotheses (or probability distributions) the events A and B should be specified in Dembski's sense. To determine this, we first consider two elements $\omega \in \Omega$ and $\lambda \in \Lambda$, such that $\lambda = \phi(\omega)$. Let h be an hypothesis that fixes a particular, but arbitrary, probability measure P_h on $\Omega \times \mathcal{K}$. Then the probability of the transformation $\omega \overset{\phi}{\mapsto} \lambda$ is of course $P_h(\omega)$. The probability that *any* randomly chosen element in Ω will be mapped onto λ is given by $Q_h^{\phi}(\lambda) = P_h(\bar{\phi}^{-1}(\lambda))$. This is easily generalized to arbitrary events and we have

$$Q_h^{\phi}(B) = P_h(\bar{\phi}^{-1}(B)). \tag{5}$$

Thus, any hypothesis h together with any mapping $\phi : \Omega \to \Lambda$ will determine a probability measure $Q_h^{\phi}(B)$ on Λ. We may there consider the original hypothesis h and the mapping ϕ to be a "partial hypothesis". However, to enable the use of Dembski's definitions, a probability measure on the larger set $\Lambda \times \mathcal{K}$ is required. There is no mathematical reason to choose any particular way to extend Eq. (5), but it seems natural to try the following assumption.

Assumption 3. *The correct way to extend Q_h^{ϕ} from a probability distribution on Λ to the larger set $\Lambda \times \mathcal{K}$ is through*

$$Q_h^{\phi}(B, K) = P_h(\bar{\phi}^{-1}(B), K). \tag{6}$$

This also seems like the extension that is most charitable to Dembski. It is not clear to me exactly which assumptions about Ω, Λ, $\mathcal{K}$ and ϕ are hidden in Eq. (6). Below I will assume that the extension is done in some unique way and index hypotheses conferring a probability distribution on $\Lambda \times \mathcal{K}$ by the original hypothesis h and the mapping ϕ.

3.1. **Is it possible to map unspecified events to specified events?** Let us now consider the first of the two conceivable ways for a deterministic process to generate specified information. Let (B, T) be h-specified information. That is, $T = \{\lambda \in \Lambda | g(\lambda) \leq \delta\} \supseteq B$ is a rejection region, there exists a $K' \subseteq \mathcal{K}$ such that $\mathcal{K}\mathrm{ID}_\Lambda(g, K')$, and $Q_h^\phi(B) = Q_h^\phi(B|K')$. Is it possible to find an $A \subseteq \Omega$, for which no valid h-specification exists and such that $\bar{\phi}(A) = B$? The answer to this question requires knowledge of the relation $\mathcal{K}\mathrm{ID}_\Omega$ and of $P_h(A, K)$. I suggest that the answer is *no*, based on the following informal appeal to the reader's intuition: If the relation $\mathcal{K}\mathrm{ID}_\Omega(f, K)$ is meaningful at all, then surely it should be true that any mathematically minded person has knowledge K that "explicitly and univocally" identifies the function $f(\omega) \equiv 0$, i.e. $\mathcal{K}\mathrm{ID}_\Omega(f, K)$. Whether it holds that $P_h(A) = P_h(A|K)$ will of course depend on h, but it seems that the case for this is at least as good as for any other function. If the reader accepts this appeal to his or her intuition, then $\Omega = \{\omega \in \Omega | f(\omega) \leq 0\}$ is a valid h-specification of A. *It then follows that every event is h-specified!*

Dembski's answer is also *no*, but he has a slightly more complicated argument in mind. He notes that the rejection function $f = g \circ \phi$ will define a set $S \subseteq \Omega$, such that $\bar{\phi}(S) = T$. Together with $\bar{\phi}(A) = B \subseteq T$, this implies that $A \subseteq S$. Dembski then seems to suggest that it is self-evident that the h-detachability of g also implies the h-detachability of $g \circ \phi$ [1, p. 153]. That is, $\mathcal{K}\mathrm{ID}_\Omega(g \circ \phi, K)$ and $P_h(A) = P_h(A|K)$, for some K. It seems unlikely to me that Dembski intended that $g \circ \phi$ is h-detachable for all possible mappings $\phi : \Omega \to \Lambda$. Rather I believe that his reasoning is most charitably interpreted as restricted to those mappings ϕ that are *known*. I speculate that part of the intended reasoning is the following informal argument: By definition, we must have knowledge that identifies g, i.e. $\mathcal{K}\mathrm{ID}_\Lambda(g, K')$ for some K'. If we also have knowledge that identifies the deterministic transformation ϕ, then this knowledge combined with K'—let's denote the combination by K—should identify the composition $g \circ \phi$, i.e. $\mathcal{K}\mathrm{ID}_\Omega(g \circ \phi, K')$. To establish that $P_h(A) = P_h(A|K)$, Dembski appears to suggest that it follows from the fact that the rejection region can be described without reference to A, and he writes that the rejection region can be "identified independently" of the event A. However, this is not necessarily the same kind of independence as the statistical independence that is required for h-detachability. We can, for instance, speak about the gasoline prices during the latest decade without ever mentioning the world's total population sizes during the same period, but that doesn't mean that gasoline prices are statistically independent of the total population sizes.

Dembski's attemped argument is of limited value due to its incompleteness and the many unclear aspects of it. It's applicability is also severly limited if I am correct that it should be interpreted as restricted to known ϕ.

3.2. **Can specifications be mapped to less probable specifications?** The second way for a deterministic process to generate specified information would be to map a specified event to an event with a smaller specification. Dembski's argument here is in fact the same as that described in the preceeding section.[2] Let us again suppose that (B,T), with $B \subseteq T \subseteq \Lambda$, is h-specified information with the rejection function g. If $\bar{\phi}(A) = B$, then Dembski thinks that one can construct a valid h-specification of A as $S = \bar{\phi}^{-1}(T) = \{\omega \in \Omega | (g \circ \phi)(\omega) \leq \delta\}$. Furthermore, by definition $Q_h^\phi(T) = P_h(\bar{\phi}^{-1}(T)) = P_h(S)$, so the probability of the h-specified information is unchanged by deterministic transformations (provided that S really is a valid h-specification).

Of course, the argument suffers from the flaws noted in the preceeding section. Dembski has not really established that there is background knowledge K such that $\mathcal{K}\mathrm{ID}(g \circ \phi, K)$ and $P_h(A) = P_h(A|K)$. If the mapping ϕ is not injective, then the latter condition may also have to hold for several events, since $\bar{\phi}(A) = B$ may not uniquely define A.

3.3. **A bizarre attempted remedy.** After having stated what I take to be the arguments described above, Dembski remarks

> "I have just argued that when a function acts to yield information, what the function acts upon has at least as much information as what the function yields. This argument, however, treats functions as mere conduits of information, and does not take seriously the possibility that functions might add information." [1, p. 154].

A truly bizarre remark. It is difficult understand what it actually means. The only flaw in Dembski's argument is that he has not established the h-detachability of $f \circ \phi$ from all events A such that $\bar{\phi}(A) = B$. Is this an admission of this flaw or did Dembski have another flaw in mind (one that I must have missed)?

Dembski then tries to repair the real or perceived flaw by considering mappings from pairs of specified information and functions to specified information. More precisely, Dembski considers a "universal composition function" U, defined by $U(i,\phi) = \bar{\phi}(i)$, where $i = (A,S) \subseteq \Omega^2$ and we take $\bar{\phi}(i) = (\bar{\phi}(A), \bar{\phi}(S))$, for notational convenience (the reader should be warned that I will change the notation slightly below). It is claimed that

> "The form of the original argument is therefore left unchanged: the information j arises by applying U (cf. f in the original argument) to the information $i,f)$ (cf. i in the original argument)." [1, p. 155].

(Note that Dembski uses f to denote the same mapping that is denoted $\bar{\phi}$ here.) It is not obvious how to interpret these statements, much less how to put them in a rigorous form. Besides from the general vagueness, it is not clear that Dembski's second argument actually improves anything. The one improvement—a big one—that can be made would be to, if possible, remedy the flaws pointed out above, but this new argument doesn't even seem to address this.

The comment that the original argument is left unchanged leads me to infer that one should primarily study how U maps events in $\Omega \times \Phi$, where $\Phi = \Lambda^\Omega = \{\phi : \Omega \to \Lambda\}$, to events in Λ, but this violates the framework provided by Dembski. To use the definitions of Dembski's key concepts, e.g. detachability, we are required to first identify a set $\mathcal{H}$ of hypotheses, each of which confers a probability measure on the sample space $\Omega \times \mathcal{K}$. If a hypothesis $h \in \mathcal{H}$ and two events in the sample space

[2]Dembski intended to treat both cases at once, whereas I have split the analysis in two cases for conceptual clarity.

$\Omega \times \mathcal{K}$ have certain properties, we may then speak of things like h-detachability, h-specified information, etc. There is no reason why hypotheses in $\mathcal{H}$ should also confer a probability measure on $\Omega \times \Phi \times \mathcal{K}$. To study this new sample space we need, in general, a new set of hypotheses, $\mathcal{H}_{\text{new}}$, conferring probability measures on *this* sample space. For a given hypothesis $h_{\text{new}} \in \mathcal{H}_{\text{new}}$ one may then speak of h_{new}-specified information, but this must not be confused with h-specified information, $h \notin \mathcal{H}_{\text{new}}$ in general, in the original sample space.

Let's examine the situation more formally. Consider a new set of hypotheses $\mathcal{H}_{\text{new}}$, each conferring an arbitrary probability measure $P_{h_{\text{new}}}$ on the sample space $\Omega \times \Phi$, where $\Phi = \{\phi : \Omega \to \Lambda\} = \Lambda^{\Omega}$. Define the mapping $U(\omega, \phi) = \phi(\omega)$ and associate two event mappings by defining

$$\bar{U}(A) = \{\phi(\omega) | (\omega, \phi) \in A\}, \quad A \subseteq \Omega \times \Phi \tag{7}$$

$$\bar{U}^{-1}(B) = \{(\omega, \phi) \in \Omega \times \Phi | \phi(\omega) \in B\}, \quad B \subseteq \Lambda. \tag{8}$$

If the outcomes in Λ are formed by applying U to outcomes in $\Omega \times \Phi$, a probability measure on Λ is fixed. The probability of the transformation $(\omega, \phi) \overset{U}{\mapsto} \phi(\omega) = \lambda$ is of course $P_{h_{\text{new}}}(\omega, \phi)$. The probability that U will map *any* element (ω, ϕ) into $B \subseteq \Lambda$ is then given by

$$Q^U_{h_{\text{new}}}(B) = P_{h_{\text{new}}}(\bar{U}^{-1}(B)), \quad B \subseteq \Lambda \tag{9}$$

This probability measure must then be extended in some way to the sample space $\Lambda \times \mathcal{K}$, so that background knowledge can be handled.

Now suppose that $(B, T) \subseteq \Lambda^2$ is h-specified information, where $h \in \mathcal{H}$ is a hypothesis in the *original* set of hypotheses, and that $\bar{\phi}(A) = B$, for some ϕ and A. The real issue is whether this can be used to construct a valid h-specification of A with probability $\leq Q^{\phi}_h(T)$, but Dembski seems to recognize that this is not always possible. Instead he seems to try to move the goalposts by considering the irrelevent question of whether one can always construct a valid h_{new}-specification of an event like $A \times \{\phi\}$ (where $h_{\text{new}} \in \mathcal{H}_{\text{new}}$ belongs to the new set of hypotheses). Perhaps Dembski thinks that $S = \bar{U}^{-1}(T)$ is always a valid h_{new}-specification of $A \times \{\phi\}$, but that would be to confuse two different kinds of specified information. If a deterministic transformation like $\omega \overset{\phi}{\mapsto} \lambda$ generates h-specified information then that means just that. Such a fact would not be affected at all by whether or not the transformation $(\omega, \phi) \overset{U}{\mapsto} \lambda$ generates h_{new}-specified information.

I challenge ID advocates, Dembski in particular, to explain in mathematical detail how the argument involving the universal composition function is supposed to work. As it is presented in [1, pp. 154–155] it is nonsense.

4. Specificational and Probabilistic Resources

Before dealing with the stochastic case it is useful to know two additional concepts defined by Dembski. The first of these is *specificational resources.* This concept is supposed to capture how many opportunities for specifications there are. The basic idea is that specified information is more remarkable if valid specifications are rare than if they are abundant. Dembski's (quasi-)formal way to deal with this idea is to introduce a "complexity measure", which measures how complicated it is for a particular user of Dembski's method to come up with a specification of an event. Let $\varphi(R)$ denote a quantitative measure of how complicated it is for a user to formulate the specification R, i.e. to find the corresponding rejection function. In [1, p. 76], Dembski asserted that the measure φ is objectively given relative to the user and "determined up to monotonic transformations". If the user of his method for inferring ID may be thought of as a universal Turing machine, then Dembski would likely take $\varphi(R)$ to be the Kolmogorov complexity of R. It is unclear how one should choose φ if the user may not be thought of as a universal Turing machine.

Definition 5 (Specificational resources)**.** *Given an event $A \subseteq \Omega$ and a rejection region R that is a valid h-specification of A for all $h \in \mathcal{H}$, the* specificational resources *are defined as the set of events $T \subseteq \Omega$ that satisfy the following conditions:*

(1) *a rejection function for T is h-detachable from A for all $h \in \mathcal{H}$,*
(2) $\max_{h\in\mathcal{H}} P_h(T) \leq \max_{h\in\mathcal{H}} P_h(R)$*,*
(3) $\varphi(T) \leq \varphi(R)$*, and*
(4) *T is not a proper subset of any other event that satisfies the previous criteria.*

The set of specificational resources will be denoted $\mathcal{S}_{\mathcal{H}}(A, R)$. [1, p. 77]

The idea is that since

$$P_h\left(\bigcup_{T\in\mathcal{S}_{\{h\}}(A,R)} T\right) \leq \sum_{T\in\mathcal{S}_{\{h\}}(A,R)} P_h(T) \leq P_h(R)|\mathcal{S}_{\{h\}}(A,R)|, \tag{10}$$

specifications with probability $\max_{h\in\mathcal{H}} P_h(R) \ll |\mathcal{S}_{\mathcal{H}}(A,R)|^{-1}$ are rare, unlikely, and remarkable.

However, even very unlikely events can become likely to occur at least once if a large number of trials occurs. To take this into account, one can multiply the specificational resources by the number of trials.

Definition 6 (Probabilistic resources)**.** *Given an event $A \subseteq \Omega$ and a rejection region R that is a valid h-specification of A for all $h \in \mathcal{H}$, the* probabilistic resources *are given by the number $\mathcal{R}_{\mathcal{H}}(A,R) = |\mathcal{S}_{\mathcal{H}}(A,R)|\cdot n$, where n is the number of trials.*

The probability that at least one event that fits a specification in $\mathcal{S}_{\mathcal{H}}(A,R)$ occurs is less than $\max_{h\in\mathcal{H}} P_h(R)\mathcal{R}_{\mathcal{H}}(A,R)$, provided that the trials are statistically independent (and that they occur according to one of the hypotheses in $\mathcal{H}$). Therefore, Dembski seems to think, an event of probability $\ll \mathcal{R}_{\mathcal{H}}(A,R)^{-1}$ should definitely be rare, unlikely, and inexplicable in terms of the hypotheses in $\mathcal{H}$. Indeed, Dembski's central claim is that specified events with such a low probability are infallible indications that the hypotheses $\mathcal{H}$ should be rejected in favour of the general ID hypothesis that some intelligent agent was somehow involved in the realization of the event (or at least some hypothesis not in $\mathcal{H}$).

There are a few problems with this reasoning, though. First, the condition that elements in $\mathcal{S}_{\mathcal{H}}(A,R)$ must be h-detachable from A with respect to all $h \in \mathcal{H}$ can seriously underestimate the opportunities for choosing specifications. For instance, suppose we have $\mathcal{H} = \{h_1, h_2\}$ and rejection functions $f_1, f_2, \ldots, f_{100}$ that are all different and identified by background knowledge. In this case it is possible that f_k is h_1-detachable from A for $1 \leq k \leq 50$ and h_2-detachable from A for $50 \leq k \leq 100$. This implies that $\mathcal{S}_{\mathcal{H}}(A,R) = \{R\}$, where R must be the rejection region defined by f_{50}, yet there are obviously more than one way to choose a specification. Dembski's error is that he considered only specifications that are valid for *all* $h \in \mathcal{H}$, rather than those that are valid for *at least one* $h \in \mathcal{H}$.

Second, there is another way in which the specificational resources underestimate the number of available specifications. Consider, for example, a random number generator that is known to generate integers between 0 and 999 with the same probability for each number. The sample space is then $\Omega = \{k \in \mathbb{N} | 0 \leq k \leq 999\}$. Let h_{ep} be the hypothesis that the outcomes are equiprobable and take $\mathcal{H} = \{h_{\mathrm{ep}}\}$. Suppose that all rejection functions of the form

$$f_k(\omega) = \begin{cases} 0 & \text{if } \lfloor \omega/5 \rfloor = k, \\ 1 & \text{otherwise,} \end{cases} \tag{11}$$

$0 \leq k \leq 199$, are h_{ep}-detachable from all events. Suppose further that no other rejection function is h_{ep}-detachable from any event. Then the rejection regions[3] $\{0,1,2,3,4\}, \{5,6,7,8,9\}, \ldots, \{995,996,997,998,999\}, \Omega$ are valid h_{ep}-specifications of any event that is a subset of them. If the event $A = \{0\}$ is observed, we then have the h_{ep}-specification $T = \{0,1,2,3,4\}$. Furthermore, $\mathcal{S}_{\mathcal{H}}(A,T) = \{T\}$. Since

$$P_{h_{\mathrm{ep}}}(T) = \frac{1}{200} \ll 1 = |\mathcal{S}_{\mathcal{H}}(A,T)|^{-1}, \tag{12}$$

one might conclude that A is a remarkable event that warrants the conclusion that h_{ep} should be rejected in favour of the general ID hypothesis. However, this would clearly be an unjustified conclusion since the same argument could be made for *any* outcome. This counter-example illustrates a deficiency in the concept of specificational resources.

While the strategy to deal with many (statistically independent) trials by multiplying the specificational resources by the number trials serves its purpose well, the concept of specificational resources does not adequately make up for the fact that the specification is chosen after the fact to fit, if possible, the observed event very tightly. The specificational resources are in a sense too tightly linked to the event that is actually observed. A more useful quantity to know would be the probability that an *outcome* is part of an event that fits an h-specification S, such that $P_h(S) \leq p$ and $\varphi(T) \leq c$, for given p, c. Let's denote this quantity by $s(p,c,h)$. That is, let

$$s_\omega(p,c,h) = \begin{cases} 1 & \text{if } \omega \text{ is contained in any event with an } h\text{-specification} \\ & T \text{ satisfying } P_h(T) \leq p \text{ and } \varphi(T) \leq c, \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

Then we have

$$s(p,c,h) = \sum_{\omega \in \Omega} P_h(\omega) s_\omega(p,c,h). \tag{14}$$

Given an observed event A with an h-specification T, quantities like

$$s(P_h(T), \varphi(T), h) \quad \text{and} \quad s(P_h(T), \infty, h)$$

are far more interesting and useful than the specificational resources $\mathcal{S}_{\mathcal{H}}(A,T)$ that Dembski wants to study. Contrary to Dembski's thinking, there is actually little basis for considering h-specifications of probability $\ll |\mathcal{S}_{\mathcal{H}}(A,T)|^{-1}$ to be remarkable, because they may well be abundant and together cover large portions of the sample space.

[3] Note that $\{0,1,2,3,4\} = \{\omega \in \Omega | f_0(\omega) \leq 0\}$, etc. and $\Omega = \{\omega \in \Omega | f_{17}(\omega) \leq 1\}$, so these events are rejection regions.

4.1. **The universal probability bound.** Let's ignore the deep conceptual flaws in the concept of probabilistic and specificational resources for the moment and consider a related concept. Determining the number of specificational resources in practice is a formidable task (assuming that the concept is well-defined), so Dembski tries make life easier by providing what is intended to be a safe, universal overestimate. To do this, he notes that the number of elementary particles in the *directly observable* universe is estimated to 10^{80}. Furthermore, he assumes that the age of the universe is less than 10^{25} seconds, or $3 \cdot 10^{17}$ years. He reasons that no transition can happen faster than the Planck time and arrives at the probabilistic resources

$$\alpha^{-1} = \frac{10^{80} \cdot 10^{25} \text{ sec}}{t_{\text{Planck}}} = 10^{150}. \tag{15}$$

At least two things should be noted. First, the universe is much bigger than the part that is directly observable. Dembski is aware of some of the current views of cosmologists and astronomers regarding this, so it is not a direct oversight [1, pp. 84–90]. However, he has seriously misunderstood the motivation behind cosmological and astronomical models. Cosmologists and astronomers are more than anything else trying to formulate descriptive models of the universe. In order to describe and explain features of our universe, a model of a universe that is bigger than the directly observable universe seems to be required (in the same sense that Lorentz invariance, at least at low energies, seems to be required of any reasonable model to describe and explain the Michelson-Morley experiments and other relativistic effects). To the best of my knowledge, no model that postulates that only the directly observable universe exists can satisfactorily account for as many astronomical and cosmological observations as our current models can account for. It appears, however, that Dembski considers cosmological models to be nothing more than desperate attempts to avoid the conclusion that God created the universe. This misunderstanding of scientists' motivation leads him to dismiss the idea that the universe is larger than the directly observable part as unfounded. In reality, it is a straightforward conclusion from the present cosmological models. For instance, it has been argued, based on the inflationary model, that the number of regions of about the size of our own directly observable region is infinite, while the number of distinct possible histories in such a region is finite [5]. Another example is the recently proposed competitor to the inflationary model that describes a cyclic universe that may be infinitely old [7]. Nonetheless, the primary motivation behind Dembski's ideas is to modify the present theories of abiogenesis and biological evolution. For this purpose it is reasonable to ignore cosmological considerations (at present there is no serious theory of abiogenesis or biological evolution that makes use of cosmological factors).

The second thing to note is that Dembski has not provided any convincing reason for interpreting α^{-1} as a safe overestimate of the number of specificational resources for even specific events in the directly observable universe. A number of concerns can be raised. E.g., how is the probability and "complexity measure" of the potential specifications taken into account, why can't more than one specification per particle arise in a time step, etc. The burden is on Dembski to show that α^{-1} really is an overestimate.[4] One can imagine a more sophisticated argument,

[4] I also note an internal inconsistency in Dembski's view of intelligence. In his framework, intelligence is usually treated as something so powerful and mystical that it cannot be analyzed

though. It has been argued that our observable universe has performed about 10^{120} operations on at most 10^{90} bits [6]. Assuming that the user's knowledge $\mathcal{K}$ is encoded by these bits and that the same knowledge can identify at most one rejection function, it is reasonable to conclude that at most 10^{90} rejection functions can be identified.[5] If only the brain of the user of Dembski's method is considered, the number will be much lower. Thus, it seems that Dembski's argument, while flawed, actually yielded a good conclusion. It seems to me that α^{-1} really is a reasonable overestimate of the specificational resources for non-astronomical purposes.

It must not be forgotten that the concept of specificational resources itself is *not* an adequate tool for determining whether a specification is remarkable or not, though. Fortunately, Dembski's *non sequitur* can be fixed if the concept of specificational resources is replaced by the degree to which specifications cover the sample space. Note that if at most N_{rf} rejection functions are "explicitly and univocally" identified by background knowledge, the degree to which the sample space is covered by h-specifications of probability $\leq p$ is bounded by

$$s(p, \infty, h) \leq pN_{\text{rf}}. \tag{16}$$

If $N_{\text{rf}} = 10^{90}$ is used, it follows that h-specifications of probability $\ll 10^{-90}$ should be rare (assuming that events really occur with the frequencies dictated by P_h).

In conclusion, the universal probability bound is an acceptable one for biological purposes, but Dembski's reasoning leading this quantity suffers from serious conceptual errors.

4.2. **CSI.** One more concept remains before it is time to investigate Dembski's argument for stochastic transformations.

Definition 7 (Complex h-Specified Information or h-CSI)**.** *Any h-specified information (A, S) with a probability $P_h(S) < \alpha$ constitutes* complex h-specified information *or* h-CSI.

For rhetorical reasons, Dembski prefers to transform probabilities of specifications by taking the negative base two logarithm and refer to the resulting number as the "complexity" (or "information") of the event.[6] The quantity α is transformed to $-\log_2(\alpha) \approx 500$ in this way, and therefore Dembski often refers to the 500 bit threshold for something to constitute h-CSI (the quantity "500 bits" must not be

rationally. Hence the privileged position that hypotheses involving intelligence occupy in the framework (such hypotheses are not accepted or rejected based on their merits or lack thereof—they are default explanations to be invoked whenever we cannot think of a specific alternative). In attempting to count specifications, however, Dembski assumes that the ability of an intelligent mind to formulate specifications is reducible to interactions of elementary particles. While he is certainly free make such a materialistic assumption, he should do consistently. In response to criticism, Dembski wrote: "We're dealing with intelligences and not with natural laws, regularities, or algorithms. It takes creativity, for instance, for a detective to see a pattern that incriminates a clever villain, and that creativity cannot be captured by 'well-defined methods.' " [2, p. 3]. Yet when it comes to counting specifications intelligence is suddenly reducible to particle physics.

[5] Or should it be $2^{10^{90}}$ rejection functions?

[6] Dembski mistakenly states that "$-\log_2 \mathbf{P}(A)$ measures the average number of bits required to specify an event with that probability" [1, p. 140]. The statement is in fact meaningless since there is nothing to average over. I believe part of the reason for this mistake is that Dembski doesn't adequately distinguish between outcomes and events. It is true that if outcomes are encoded by bit strings, then $-\log_2(p)$ is a good guess for how many bits that are needed to encode an outcome of probability p in the encoding that minimizes the average code word length. This, however, cannot be directly generalized to events.

confused with 500 binary digits—in this case the unit serves no purpose except to indicate that logarithms were taken to base two).

5. The Law of Conservation of Information: The Stochastic Case

When it comes to determining whether or not stochastic transformations can generate h-specified information, Dembski has little to add to his argument for the deterministic case. First, he admits that stochastic processes can generate some h-specified information, but asserts that h-specified information which is improbable enough to constitute h-CSI cannot be generated. To support this he then notes that any stochastic transformation may be viewed as a two-step process without any loss of generality. First, a *deterministic transformation* is chosen randomly according to a probability function ρ. Then this mapping is applied to the outcome $\omega \in \Omega$ in question. But Dembski thinks that he has established that no h-specified information can be generated by *any* deterministic transformation, and therefore simply refers back his argument with the "universal composition function" for the deterministic case. After all, the argument is, if no deterministic transformation can generate h-specified information, then it won't matter which deterministic mapping is chosen at the first step. This argument, however, contradicts Dembski's admission that stochastic transformations can generate a little h-specified information. If Dembski's argument were correct, then stochastic transformations would never generate any h-specified information at all.

6. Conclusion

In the preceeding sections the argument used by Dembski to establish the so-called Law of Conservation of Information has been investigated mathematically. I have intentionally been forthcoming with Dembski's philosophical assumptions about the nature of knowledge and the privileged status of the general ID hypothesis in order to focus on the mathematical parts of the argument. From his own words it seems that he considers his argument to have a status close to a mathematical proof. Indeed, he wrote:

> "In this section I will present an in-principle mathematical argument for why natural causes are incapable of generating complex specified information." [1, p. 150]
>
> "Justifying the claim that natural causes cannot generate complex specified information is technically demanding. Before justifying this claim mathematically, let me therefore try to spell out in plain English why natural causes are not up to the task of generating CSI." [1, p. 150]

Such statements indicate that Dembski considers his argument to be if not a mathematical proof, so at least a pretty rigorous mathematical argument. However, a mathematical investigation of Dembski's quasi-formalization reveals that the argument is far from rigorous. It suffers from several conceptual flaws, some of which are reparable and some of which seem fatal. The main flaws are:

(1) In the original argument for the deterministic case, Dembski assumes that if the rejection function f on Λ is detachable, then so is the rejection function $f \circ \phi$ on Ω. But this is by no means established.

(2) The argument involving the "universal composition function" confuses two different kinds of specified information.
(3) It is possible for a number of improbable specifications to together cover a very large portion of the sample space. Given an event A and a h-specification T, there is nothing that guarantees that $s(P_h(T), \varphi(T), h)$ (as defined by Eq. (14)) is low just because $P_h(T) \ll |\mathcal{S}_{\mathcal{H}}(A,T)|^{-1}$. The concept of specificational resources therefore does not adequately serve its purpose and it should be replaced by the quantity defined by Eq. (14).
(4) Dembski admits that stochastic transformations can generate some specified information (but not enough to constitute CSI). However, if his own argument were correct, it would be impossible for stochastic transformations to generate any specified information at all.

The Law of Conservation of Information must therefore be regarded as mathematically unsubstantiated. Regardless of the non-mathematical virtues, or lack thereof, of Dembski's argument, ambiguities and equivocation makes it entirely inadequate as a mathematical argument.

6.1. **Acknowledgement.** I thank Mark Perakh for useful comments on earlier drafts.

References

[1] Dembski, William A. (2002) "No Free Lunch", *Rowman & Littlefield*
[2] Dembski W. (2002) "The Fantasy Life of Richard Wein: A Response to a Response", http://www.iscid.org/papers/Dembski_WeinsFantasy_060702.pdf
[3] Elsberry W. & Shallit J. (2002) "Information Theory, Evolutionary Computation, and Dembski's 'Complex Specified Information' ", manuscript submitted for publication
[4] Falk A. (2001) "Can We Be Darwinians and Religious?", *Quarterly Review of Biology*, **76**(1) : 47–52 (book rewiew)
[5] Garriga J. & Vilenkin A. (2001) "Many worlds in one", http://xxx.lanl.gov/abs/gr-qc/0102010
[6] Lloyd S. (2002) "Computational Capacity of the Universe", *Physical Review Letters*, **88**(23) : 237901
[7] Steinhardt P. & Turok N. (2002) "A Cyclic Model of the Universe", *Science*, **296**(5572) : 1436–1439
[8] Wein R. (2002) "Not a Free Lunch But a Box of Chocolates", http://www.talkorigins.org/design/faqs/nfl/