

# On a Wrong Remark in a Paper of Robert J. Marks II and William A Dembski

D. Eben

March 13, 2012

## Abstract

In their 2010 paper *The Search for a Search: Measuring the Information Cost of Higher Level Search* [1], the authors William A. Dembski and Robert J. Marks II present as one of two results their so-called *Horizontal No Free Lunch Theorem*. One of the consequences of this theorem is their remark: *If no information about a search exists, so that the underlying measure is uniform, then, on average, any other assumed measure will result in negative active information, thereby rendering the search performance worse than random search.* This is quite surprising, as one would expect in the tradition of the *No Free Lunch* theorem that the performances are equally good (or bad). Using only very basic elements of probability theory, this essay shows that their remark is wrong - as is their theorem.

## 1 Definitions

The situation is quite simple: There is a set  $\Omega_1$  - the search space - with  $N$  elements, so w.l.o.g.  $\Omega = \{1, \dots, N\}$ . Let  $T_1 \subset \Omega_1$  denote the target. At first assume that  $T_1$  consists of a single element. To find this element  $T_1$  you have  $K$  guesses and you aren't allowed to guess the same element twice. Such a sequence of  $K$  different guesses can be described as an ordered  $K$ -tuple  $(n_1, n_2, \dots, n_K)$  of  $K$  pairwise different numbers and is called a *search*. The  $\frac{N!}{(N-K)!}$  different searches build the space  $\Omega_K$ .

A *search strategy* is a probability measure  $\mathbf{Q}$  on  $\Omega_K$ . Such a measure can be given by assigning a probability  $\mathbf{Q}((n_1, n_2, \dots, n_K)) = q_{(n_1, n_2, \dots, n_K)}$  for each of the (pure) strategies  $(n_1, n_2, \dots, n_K)$  in  $\Omega_K$ :  $q_{(n_1, n_2, \dots, n_K)}$  is the probability to choose the strategy  $(n_1, n_2, \dots, n_K)$  from  $\Omega_K$ . So obviously,  $q_{(n_1, n_2, \dots, n_K)} \geq 0 \forall (n_1, n_2, \dots, n_K) \in \Omega_K$  and  $\sum_{(n_1, n_2, \dots, n_K) \in \Omega_K} q_{(n_1, n_2, \dots, n_K)} = 1$ .

*Random search* (i.e., the *random search strategy*) correspond to the uniform distribution  $\mathbf{U}$  on  $\Omega_K$ , i.e.  $\mathbf{U}((n_1, n_2, \dots, n_K)) = \frac{(N-K)!}{N!}$ .

Any other search strategy is called an *assisted search*. This shows the small scope of the concept of an assisted search in the papers of Dembski and Marks: it is not possible to encode the influence of an oracle or a fitness function into the measure  $\mathbf{Q}$ , so most examples presented by the authors (like *partitioned search*[2] or *easter egg hunt*[1]) are not covered by this model.

The *performance* of a search strategy  $\mathcal{P}(\mathbf{Q}, \delta_{T_1})$  for finding an element  $T_1 \in \Omega_1$  is defined as the probability to name  $T_1$  in one of the  $K$  guesses of the search.

## 2 Some calculations

Using the characteristic function

$$\chi_{\{n_1, n_2, \dots, n_K\}}(T_1) = \begin{cases} 1 & \text{if } T_1 \in \{n_1, n_2, \dots, n_K\} \\ 0 & \text{otherwise} \end{cases}$$

this probability can be written as:

$$\mathcal{P}(\mathbf{Q}, \delta_{T_1}) = \sum_{(n_1, n_2, \dots, n_K) \in \Omega_K} q_{(n_1, n_2, \dots, n_K)} \cdot \chi_{\{n_1, n_2, \dots, n_K\}}(T_1)$$

The characteristic function is independent of the order of the elements  $n_1, n_2, \dots, n_K$ . So the expression can be streamlined using instead of the space of *ordered n-tuples*  $\Omega_K$  the space  $\mathfrak{P}_K(\Omega_1)$  of subsets  $\{n_1, n_2, \dots, n_K\}$  with  $K$

elements of  $Q_1$ . This space has  $\binom{N}{K}$  elements.  $\mathbf{Q}$  on  $\Omega_K$  gives instantly rise to a measure on  $\mathfrak{P}_K(\Omega_1)$  - again called  $\mathbf{Q}$  - by setting:

$$\mathbf{Q}(\{n_1, n_2, \dots, n_K\}) = \sum_{\sigma \in S_K} q(\sigma_1(n_1), \sigma_2(n_2), \dots, \sigma_K(n_K)).$$

Here,  $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_K)$  are the  $K!$  elements of the group of permutations of  $K$  elements,  $S_K$ . Thus, the probability to search successfully for  $T_1$  is

$$\mathcal{P}(\mathbf{Q}, \delta_{T_1}) = \sum_{\{n_1, n_2, \dots, n_K\} \in \mathfrak{P}_K(\Omega_1)} \mathbf{Q}(\{n_1, n_2, \dots, n_K\}) \cdot \chi_{\{n_1, n_2, \dots, n_K\}}(T_1).$$

But how does this search perform on average - assuming *that the underlying measure is uniform*? The *underlying measure* gives us the probability with which a target element is chosen: if we have a uniform underlying measure  $\mathbf{U}_\Omega$ , each element  $1, 2, \dots, N$  is a the target with a probability of  $\frac{1}{N}$ . The average performance is given by:

$$\begin{aligned} \mathcal{P}(\mathbf{Q}, \mathbf{U}_\Omega) &= \sum_{l=1}^N \frac{1}{N} \sum_{\{n_1, n_2, \dots, n_K\} \in \mathfrak{P}_K(\Omega_1)} \mathbf{Q}(\{n_1, n_2, \dots, n_K\}) \cdot \chi_{\{n_1, n_2, \dots, n_K\}}(l) \\ &= \frac{1}{N} \sum_{\{n_1, n_2, \dots, n_K\} \in \mathfrak{P}_K(\Omega_1)} \mathbf{Q}(\{n_1, n_2, \dots, n_K\}) \cdot \left( \sum_{l=1}^N \chi_{\{n_1, n_2, \dots, n_K\}}(l) \right) \\ &= \frac{1}{N} \sum_{\{n_1, n_2, \dots, n_K\} \in \mathfrak{P}_K(\Omega_1)} \mathbf{Q}(\{n_1, n_2, \dots, n_K\}) \cdot K \\ &= \frac{K}{N} \sum_{\{n_1, n_2, \dots, n_K\} \in \mathfrak{P}_K(\Omega_1)} \mathbf{Q}(\{n_1, n_2, \dots, n_K\}) \\ &= \frac{K}{N}. \end{aligned}$$

This holds for every measure  $\mathbf{Q}$  on the space of the searches, especially for the uniform measure, leading to the

**Special Remark:** IF NO INFORMATION ABOUT A SEARCH FOR A SINGLE ELEMENT EXISTS, SO THAT THE UNDERLYING MEASURE IS UNIFORM, THEN, ON AVERAGE, ANY OTHER ASSUMED MEASURE WILL RESULT IN THE SAME SEARCH PERFORMANCE AS RANDOM SEARCH.

But perhaps this result depends on the special condition of an elementary target? What happens when  $T_m = \{t_1, t_2, \dots, t_m\} \in \mathfrak{P}_m(\Omega_1)$ ? Then a search is seen as a success if it identifies at least one element of the target  $T_m$ , thus

$$\mathcal{P}(\mathbf{Q}, \delta_{T_1}) = \sum_{\{n_1, n_2, \dots, n_K\} \in \mathfrak{P}_K(\Omega_1)} q_{\{n_1, n_2, \dots, n_K\}} \cdot \max_{t \in T_1} \{\chi_{\{n_1, n_2, \dots, n_K\}}(t)\}.$$

Calculating the average over all elements of  $\mathfrak{P}_m(\Omega_1)$  results in:

$$\begin{aligned} \mathcal{P}(\mathbf{Q}, \mathbf{U}_{\mathfrak{P}_m(\Omega_1)}) &= \frac{1}{\binom{N}{m}} \sum_{T \in \mathfrak{P}_m(\Omega_1)} \sum_{\{n_1, n_2, \dots, n_K\} \in \mathfrak{P}_K(\Omega_1)} q_{\{n_1, n_2, \dots, n_K\}} \cdot \max_{t \in T} \{\chi_{\{n_1, n_2, \dots, n_K\}}(t)\} \\ &= \frac{1}{\binom{N}{m}} \sum_{\{n_1, n_2, \dots, n_K\} \in \mathfrak{P}_K(\Omega_1)} q_{\{n_1, n_2, \dots, n_K\}} \sum_{T \in \mathfrak{P}_m(\Omega_1)} \max_{t \in T} \{\chi_{\{n_1, n_2, \dots, n_K\}}(t)\} \\ &= \frac{1}{\binom{N}{m}} \sum_{\{n_1, n_2, \dots, n_K\} \in \mathfrak{P}_K(\Omega_1)} q_{\{n_1, n_2, \dots, n_K\}} \cdot \left( \binom{N}{m} - \binom{N-K}{m} \right) \\ &= \frac{\binom{N}{m} - \binom{N-K}{m}}{\binom{N}{m}} \sum_{\{n_1, n_2, \dots, n_K\} \in \mathfrak{P}_K(\Omega_1)} q_{\{n_1, n_2, \dots, n_K\}} \\ &= \frac{\binom{N}{m} - \binom{N-K}{m}}{\binom{N}{m}} = 1 - \frac{(N-K)(N-K-1) \dots (N-K-m+1)}{N(N-1) \dots (N-m+1)}. \end{aligned}$$

Again the probability of a successful search - and by definition the performance - is on average independent of the measure used on the space of the searches! (Note that the expression becomes  $\frac{K}{N}$  for  $m = 1$  and 1 for  $N - K < m$ ).

The result of this calculation is the

**Less Special Remark:** IF NO INFORMATION ABOUT A SEARCH FOR A TARGET EXISTS OTHER THAN THE NUMBER OF ELEMENTS THE TARGET EXISTS FROM, SO THAT THE UNDERLYING MEASURE IS UNIFORM, THEN, ON AVERAGE, ANY OTHER ASSUMED MEASURE WILL RESULT IN THE SAME SEARCH PERFORMANCE AS RANDOM SEARCH.

As the performance of the random search and any other search doesn't differ on any subset of targets of fixed cardinality, the next remark is elementary:

**General Remark:** IF NO INFORMATION ABOUT A SEARCH EXISTS, SO THAT THE UNDERLYING MEASURE IS UNIFORM, THEN, ON AVERAGE, ANY OTHER ASSUMED MEASURE WILL RESULT IN THE SAME SEARCH PERFORMANCE AS RANDOM SEARCH.

Obviously this **General Remark** negates the remark of Dembski and Marks:

IF NO INFORMATION ABOUT A SEARCH EXISTS, SO THAT THE UNDERLYING MEASURE IS UNIFORM, THEN, ON AVERAGE, ANY OTHER ASSUMED MEASURE WILL RESULT IN NEGATIVE ACTIVE INFORMATION, THEREBY RENDERING THE SEARCH PERFORMANCE WORSE THAN RANDOM SEARCH.

### 3 Consequences for the *Horizontal No Free Lunch* Theorem

There is much critique on this theorem out there, e.g., "A Note Regarding "Active Entropy""[4]. The previous paragraphs add to this: Dembski's and Marks's *remark* is a direct consequence of the theorem, and as the remark is wrong, the theorem can't be right, neither.

How is it possible for two mathematicians to err on such a crucial point? Perhaps they were seduced by the elegance of their argument: the calculations in the sections above are simple and a little bit clumsy. Dembski and Marks smartly augmented the original search space  $\Omega_1$  and identified it with the space of pure search strategies  $\Omega_K$ . For each original target  $T_1$  a corresponding target  $\tilde{T}_1$  can be found in  $\Omega_K$ , such that a simple search  $(n_1, \dots, n_K)$  is successful when it is an element of  $\tilde{T}_1$ . So the problem of searching by  $K$  guesses on  $\Omega_1$  is reduced to a simple guess on  $\Omega_K$ . The problem which they overlooked: the corresponding targets aren't nice subsets. Especially if  $K > 1$ , then the corresponding sets for two different targets in the original space have always a non-empty intersection. Nevertheless, Dembski and Marks use an *exhaustive partition* of the augmented space in their *proof* of the *Horizontal No Free Lunch* theorem, but there is no such partition into subsets which makes sense in the framework of their argument.

### 4 Conclusion

Dembski and Marks have introduced their concept of *Active Information* using a great number of examples[2][3]. Their model for an assisted search works only for one or two of these examples. The theoretical results presented in *The Search for a Search*[1] seem not to work at all. That leaves the *Active Information* as being a new performance measure, but it can't be seen how it improves the understanding of the concept of information in connection with search algorithms.

### References

- [1] William A. Dembski and Robert J. Marks II, "The Search for a Search: Measuring the Information Cost of Higher Level Search," Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol.14, No.5, 2010, pp. 475-486.
- [2] William A. Dembski and Robert J. Marks II "Conservation of Information in Search: Measuring the Cost of Success" IEEE Transactions on Systems, Man and Cybernetics A, Systems & Humans, Vol.5, No.5, September 2009, pp.1051-1061
- [3] Winston Ewert, George Montanez, William A. Dembski, Robert J. Marks II, "Efficient Per Query Information Extraction from a Hamming Oracle," Proceedings of the 42nd Meeting of the Southeastern Symposium on System Theory, IEEE, University of Texas at Tyler, March 7-9, 2010, pp.290-297.
- [4] anonymous, "A Note Regarding "Active Entropy""