

# Sophistication as Randomness Deficiency

Francisco Mota<sup>1,2</sup>, Scott Aaronson<sup>4</sup>, Luís Antunes<sup>1,2</sup>, and André Souto<sup>1,3</sup>

<sup>1</sup> Security and Quantum Information Group at Instituto de Telecomunicações  
fmota@fmota.eu, lfa@dcc.fc.up.pt, asouto@math.ist.utl.pt

<sup>2</sup> Departamento de Ciência de Computadores at FCUP

<sup>3</sup> Departamento de Matemática at IST at UTL

<sup>4</sup> Computer Science and Artificial Intelligence Lab at MIT  
aaronson@csail.mit.edu

**Abstract.** The sophistication of a string measures how much structural information it contains. We introduce *naïve sophistication*, a variant of sophistication based on randomness deficiency. Naïve sophistication measures the minimum number of bits needed to specify a set in which the string is a typical element. Thanks to Vereshchagin and Vítányi, we know that sophistication and naïve sophistication are equivalent up to low order terms. We use this to relate sophistication to lossy compression, and to derive an alternative formulation for busy beaver computational depth.

## 1 Introduction

Kolmogorov complexity measures the amount of information intrinsic in an object by measuring how much an object can be compressed. For example, the string  $0^{1000}$  (that is, 0 repeated 1000 times) can be compressed into only a few bits and therefore has very little information. On the other hand, a random string cannot easily be compressed and therefore has a lot of information.

Different methods of compression will yield different complexity measures. Fortunately, there is a compression scheme that yields an optimal complexity measure. The trick is to describe a string  $x$  as a pair  $\langle p, d \rangle$ , where  $p$  is a program in some prefix-free Turing-complete language, and the program  $p$  generates the string  $x$  when given the string  $d$  as input. Thus we define the Kolmogorov complexity of  $x$  as the length of the shortest description of  $x$  under this universal compression scheme:

$$K(x) = \min_{p,d} \{ |p| + |d| : \langle p, d \rangle \text{ is a description of } x \}.$$

We say that a string  $x$  is *incompressible* if  $K(x) \geq |x| - O(1)$ . Incompressible strings are indistinguishable from randomly generated strings, so we equate the two: a random string is an incompressible string. Furthermore, we know that if  $\langle p, d \rangle$  is an optimal two-part description for  $x$ , then  $d$  is incompressible. Thus we say that  $p$  represents the *structural information* of  $x$  and  $d$  represents the *random information* of  $x$ .

This sets the stage for the study of sophistication. Sophistication measures the amount of structural information in a string. We look at all short descriptions of a string and minimize over the structural information in those descriptions. An equivalent way to look at sophistication is to model a string by a finite set that contains it. Some sets are better models than others. Sophistication is then the minimum complexity of a good model for that string.

For example, the set of all strings of length  $n$  is a good way to model completely random strings of length  $n$ , so a completely random string has very low sophistication. On the other hand, the set  $\{x\}$  is always a good model for the string  $x$ , so a string of very low complexity also has very low sophistication.

One of the characteristics of a good model for  $x$  is that  $x$  must be a typical element of that model, in the sense that  $x$  is generally indistinguishable from an element of the model taken at random. Randomness deficiency addresses this question of typicality. Randomness deficiency is used to measure how typical a string is with respect to a finite set that contains it.

The contributions of this paper are:

- We introduce a sophistication measure based on randomness deficiency, *naive sophistication* (§3). Vereshchagin and Vítányi [1] showed that naive sophistication is equivalent to sophistication up to low order terms (§4).
- We relate naive sophistication to the limits of lossy compression, in the sense that naive sophistication measures how many bits are needed to represent a consistent model of the string. With a consistent model, we can query the properties of the string with a very low false positive rate (§5).
- We compare naive sophistication to computational depth (§6). By using naive sophistication, we establish an alternative definition for busy beaver computational depth.

## 2 Preliminaries

In this section we present formal definitions for Kolmogorov complexity, randomness deficiency, discrepancy, and sophistication. What follows is a brief summary of the theory of Kolmogorov complexity. For more details, we suggest the reading of [2].

**Definition 1 (Kolmogorov Complexity).** *Let  $T$  be a prefix-free Turing machine. We define the conditional Kolmogorov complexity of  $x$  given  $y$  relative to  $T$  as the length of the shortest program that outputs  $x$  when given  $y$  as an additional input:*

$$K_T(x|y) = \min_p \{ |p| : T(p, y) = x \}.$$

*There is a prefix-free Turing machine  $U$  which yields an optimal prefix-free complexity measure. For any prefix-free Turing machine  $T$ , there is a constant  $c_T$*

such that  $K_U(x|y) \leq K_T(x|y) + c_T$ . Because  $U$  is optimal, we drop the subscript, and this is what we call the Kolmogorov complexity of  $x$  given  $y$ :

$$K(x|y) = K_U(x|y) = \min_p \{ |p| : U(p, y) = x \}.$$

We also use special notation for the case where the condition  $y$  is the empty string  $\varepsilon$ . This is equivalent to the definition of  $K(x)$  presented in the introduction:

$$K(x) = K(x|\varepsilon) = \min_p \{ |p| : U(p, \varepsilon) = x \}.$$

Kolmogorov complexity measures how much information we need to describe a string. We can also extend this definition to finite sets of strings, by encoding the set  $S = \{x_1, x_2, \dots, x_k\}$  as a string  $[S] = [k][x_1][x_2] \cdots [x_k]$  where the sequence  $x_1, \dots, x_k$  enumerates all elements of  $S$  in lexicographic order,  $[k]$  is a prefix-free encoding<sup>1</sup> of the integer  $k$ , and  $[x_i]$  is a prefix-free encoding<sup>2</sup> of the string  $x_i$ . The complexity of a finite set  $S$  is then

$$K(S) = K([S]) = K([k][x_1][x_2] \cdots [x_k]).$$

Likewise, we can use a finite set  $S$  as the condition in conditional Kolmogorov complexity. We define  $K(x|S) = K(x|[S])$ . Given such a description of a set, we would need at most  $\log |S|$  bits to describe any element from the set. That is, if  $x \in S$ , then  $K(x|S) \leq \log |S| + O(1)$ . By a counting argument, most elements of  $S$  also satisfy  $K(x|S) \geq \log |S| - O(1)$ , but there might be a shorter way to describe  $x$  given  $S$ , if  $x$  is atypical.

**Definition 2 (Randomness Deficiency).** *The randomness deficiency of  $x$  with respect to a finite set  $S$  containing  $x$  is defined as*

$$\delta(x|S) = \log |S| - K(x|S).$$

*An element  $x$  of  $S$  with low randomness deficiency is said to be typical, for the best way to describe  $x$  using  $S$  is to pinpoint its exact location in the set. In this paper we adopt the convention that if  $x$  is not an element of  $S$ , then  $\delta(x|S) = \infty$ .*

We could also use  $S$  to describe any  $x \in S$  unconditionally, by first describing  $S$  and then pinpointing  $x$ 's location in it:  $K(x) \leq K(S) + \log |S| + O(1)$ . However, the set  $S$  might be a very poor model of  $x$ , resulting in a large gap between  $K(x)$  and  $K(S) + \log |S|$ . We introduce *discrepancy* as a new notation to talk about this gap.

**Definition 3 (Discrepancy).** *The discrepancy of  $x$  with respect to a finite set  $S$  containing  $x$  is the additional cost of using  $S$  to describe  $x$ . It is defined as*

$$\Delta(x|S) = \log |S| - K(x) + K(S).$$

*By convention, if  $x$  is not an element of  $S$ , then  $\Delta(x|S) = \infty$ .*

<sup>1</sup> For example, let  $[k] = 0^k 1$ , for any natural number  $k$ .

<sup>2</sup> For example, let  $[x] = [|x|]x = 0^{|x|} 1x$ , for any string  $x$ .

Randomness deficiency measures how far  $x$  is from being a typical element of  $S$ , and discrepancy measures how far  $S$  is from being a good model of  $x$ . They are two sides of the same coin, and they are simply related. It is easy to see that  $\delta(x|S) \leq \Delta(x|S) + O(1)$  for any  $x$  and  $S$ . The following lemma tells us exactly how far apart they are, up to a logarithmic term.

**Lemma 1.** *For any finite set of strings  $S$  and any string  $x \in S$ , we have*

$$\Delta(x|S) = \delta(x|S) + K(S|x)$$

*up to a logarithmic additive term in  $K(S)$  and  $K(x)$ .*

*Proof.* By the symmetry of algorithmic information [3], we know that  $K(S) - K(S|x) = K(x) - K(x|S)$  up to a logarithmic additive term in  $K(S)$  and  $K(x)$ . Rearranging the terms and adding  $\log |S|$  to both sides gives us the desired approximate equality.  $\square$

## 2.1 Sophistication and Coarse Sophistication

Sophistication, as defined by Koppel [4], measures the amount of structural information contained in a string. According to Vitányi [5], this is equivalent to measuring the complexity of a good model for that string, up to low order terms.

**Definition 4 (Sophistication).** *The sophistication of  $x$  is the complexity of the simplest model of  $x$  with limited discrepancy:*

$$\text{soph}_c(x) = \min_S \{ K(S) : \Delta(x|S) \leq c \}.$$

*The significance level  $c$  tells us how much discrepancy  $S$  is allowed to have. We say that  $S$  is a witness to  $\text{soph}_c(x)$  if  $K(S) \leq \text{soph}_c(x)$  and  $\Delta(x|S) \leq c$ .*

Antunes and Fortnow [6] defined a variant of sophistication, called *coarse sophistication*, that gets rid of the significance level by incorporating it into the minimization.

**Definition 5 (Coarse Sophistication).** *The coarse sophistication of a string  $x$  minimizes both the complexity of the model and its discrepancy:*

$$\text{csoph}(x) = \min_S \{ K(S) + \Delta(x|S) \}.$$

*Once again, this definition is equivalent to the definition based on structural information, given in [6], up to a logarithmic additive term. We say that  $S$  is a witness to  $\text{csoph}(x)$  if  $K(S) + \Delta(x|S) \leq \text{csoph}(x)$ .*

What follows is an alternative definition for coarse sophistication.

**Lemma 2.** *We have:*

$$\text{csoph}(x) = \min_c \{ \text{soph}_c(x) + c \}.$$

*Proof*

( $\leq$ ) For any  $c$ , let  $S$  be a witness to  $\text{soph}_c(x)$ . By definition,  $\Delta(x|S) \leq c$ . Therefore,

$$\text{csoph}(x) \leq K(S) + \Delta(x|S) \leq \text{soph}_c(x) + c.$$

Since this is true for any  $c$ , we have  $\text{csoph}(x) \leq \min_c \{ \text{soph}_c(x) + c \}$ .

( $\geq$ ) Let  $S$  be a witness to  $\text{csoph}(x)$  and let  $d = \Delta(x|S)$ . Therefore,

$$\begin{aligned} \min_c \{ \text{soph}_c(x) + c \} &\leq \text{soph}_d(x) + d \\ &\leq K(S) + \Delta(x|S) \\ &= \text{csoph}(x). \end{aligned}$$

□

### 3 Naive Sophistication

We now define a sophistication measure based on randomness deficiency.<sup>3</sup>

**Definition 6 (Naive Sophistication).** *The naive sophistication of  $x$  is the complexity of the simplest set in which  $x$  is a typical element:*

$$\text{nsoph}_c(x) = \min_S \{ K(S) : \delta(x|S) \leq c \}.$$

*The significance level  $c$  tells us how atypical  $x$  is allowed to be. We say that  $S$  is a witness to  $\text{nsoph}_c(x)$  if  $K(S) \leq \text{nsoph}_c(x)$  and  $\delta(x|S) \leq c$ .*

**Definition 7 (Naive Coarse Sophistication).** *Naive coarse sophistication gets rid of the significance level, by minimizing over both the complexity of the set and the resulting randomness deficiency:*

$$\text{ncsoph}(x) = \min_S \{ K(S) + \delta(x|S) \}.$$

*We say that  $S$  is a witness to  $\text{ncsoph}(x)$  if  $K(S) + \delta(x|S) \leq \text{ncsoph}(x)$ .*

Naive coarse sophistication is the naive counterpart to coarse sophistication. Lemma 2 also applies to naive coarse sophistication. In other words, the following equation is an equivalent definition for naive coarse sophistication:

$$\text{ncsoph}(x) = \min_c \{ \text{nsoph}_c(x) + c \}.$$

### 4 Comparing Sophistication Measures

In this section, we show that naive sophistication is equivalent to sophistication up to low order terms, based on previous work by Vereshchagin and Vítányi [1]. These equations and inequalities hold up to a logarithmic additive term in  $|x|$ :

$$\begin{aligned} \text{soph}_{c+O(\log |x|)}(x) &\leq \text{nsoph}_{c+O(1)}(x) \leq \text{soph}_c(x), \\ \text{csoph}(x) &= \text{ncsoph}(x). \end{aligned}$$

---

<sup>3</sup> Aaronson first introduced naive sophistication in his MathOverflow question [7].

**Theorem 1.** *Naive sophistication is a lower bound for sophistication:*

$$\begin{aligned}\text{nsoph}_{c+O(1)}(x) &\leq \text{soph}_c(x), \\ \text{ncsoph}(x) &\leq \text{csoph}(x) + O(1).\end{aligned}$$

*Proof.* This is a consequence of  $\delta(x|S) \leq \Delta(x|S) + O(1)$  for any  $x$  and  $S$ .  $\square$

**Lemma 3 (Lemma A.4 of [1]).** *For any finite set  $A$  containing  $x$  with  $K(A) + \log |A| \leq O(|x|)$ , then there is a finite set  $S$  containing  $x$  with  $K(S) \leq K(A) - K(A|x) + O(\log |x|)$  and  $\log |S| \leq \log |A|$ .*

**Theorem 2.** *Naive sophistication is an upper bound for sophistication:*

$$\begin{aligned}\text{soph}_{c+O(\log |x|)}(x) &\leq \text{nsoph}_c(x) + O(\log |x|), \\ \text{csoph}(x) &\leq \text{ncsoph}(x) + O(\log |x|).\end{aligned}$$

*Proof.* Let  $A$  witness  $\text{nsoph}_c(x)$ . By Lemma 3, there is a set  $S$  with  $K(S) \leq K(A) - K(A|x) + O(\log |x|)$  and  $\log |S|$ . Therefore,

$$\begin{aligned}\Delta(x|S) &= K(S) + \log |S| - K(x) \\ &\leq (K(A) - K(A|x) + O(\log |x|)) + \log |A| - K(x) \\ &\leq \Delta(x|A) - K(A|x) + O(\log |x|) \\ &\leq \delta(x|A) + O(\log |x|) && \text{(by Lemma 1)} \\ &\leq c + O(\log |x|).\end{aligned}$$

Therefore  $\text{soph}_{c+O(\log |x|)}(x) \leq K(S) \leq \text{nsoph}_c(x) + O(\log |x|)$ . The  $\text{csoph}(x)$  upper bound follows by Lemma 2.  $\square$

We can also use naive sophistication to create an upper bound for sophistication that uses a  $O(1)$  additive term in the significance level, but it is significantly weaker.

**Theorem 3 (Upper bound for sophistication with constant overhead)**

$$\begin{aligned}\text{soph}_{\text{nsoph}_c(x)+c+O(1)}(x) &\leq \text{nsoph}_c(x), \\ \text{csoph}(x) &\leq 2 \cdot \text{nsoph}_c(x) + c + O(1).\end{aligned}$$

*Proof* Let  $S$  be a witness to  $\text{nsoph}_c(x)$ . Notice that  $\log |S| \leq K(x|S) + c$ , that  $K(S) = \text{nsoph}_c(x)$ , and that  $K(x|S) \leq K(x) + O(1)$ . We have,

$$\begin{aligned}\Delta(x|S) &\leq K(S) + \log |S| - K(x) \\ &\leq K(S) + K(x|S) + c - K(x) \\ &\leq K(S) + c + O(1) \\ &= \text{nsoph}_c(x) + c + O(1).\end{aligned}$$

The  $\text{csoph}(x)$  upper bound follows by Lemma 2.  $\square$

## 5 Relation to Lossy Compression

Naive sophistication measures the limits of lossy compression. This is true in the sense that we need only a witness to  $\text{nsoph}_c(x)$  to be able to query properties of  $x$  without false positives, and we need at least  $\text{nsoph}_{c+O(1)}(x)$  bits to describe any such model of  $x$ . The connection between lossy compression and randomness deficiency was established in [1]. We are elaborating on that connection.

Let us say that a set  $S$  is  $(c, x)$ -consistent if and only if, for all properties  $P \subseteq \Sigma^*$ , if  $P$  occurs with high probability in  $S$ , then  $x \in P$ . Formally,  $S$  is  $(c, x)$ -consistent if for all properties  $P$ ,

$$\Pr_S(P) \geq 1 - 2^{-c-K(P|S)} \quad \text{implies} \quad x \in P.$$

**Theorem 4.** *Consistency is equivalent to randomness deficiency, up to a constant  $d$ :*

1. A set  $S$  is  $(c, x)$ -consistent if  $\delta(x|S) \leq c - d$ .
2. A set  $S$  is  $(c, x)$ -consistent only if  $\delta(x|S) \leq c + d$ .

*Proof.*

1. Let  $S$  satisfy  $\delta(x|S) \leq c - d$ . We will show that  $S$  is  $(c, x)$ -consistent. Let  $P \subseteq \Sigma^*$  be a property that occurs with high probability in  $S$ . That is,  $\Pr_S(P) \geq 1 - 2^{-c-K(P|S)}$ . Let  $Q = S \setminus P$ . Then  $|Q| < |S| \cdot 2^{-c-K(P|S)}$ . If  $x \in Q$ , we have

$$\begin{aligned} K(x|S) &\leq K(Q|S) + \log |Q| \\ &< (K(P|S) + O(1)) + (\log |S| - c - K(P|S)) \\ &\leq \log |S| - c + O(1). \end{aligned}$$

This implies  $\delta(x|S) > c - O(1)$ , which is a contradiction for large enough values of  $d$ . Therefore  $x \in P$ . This holds for all properties  $P$  that occur with high probability in  $S$ , so  $S$  is  $(c, x)$ -consistent.

2. We know that for any property  $P$  with  $\Pr_S(P) \geq 1 - 2^{-c-K(P|S)}$ , we have  $x \in P$ . By way of contradiction, let us assume that  $\delta(x|S) > c + d$ . That is,  $K(x|S) < \log |S| - c - d$ . Let  $P = \{ y : y \neq x \}$ . Note that  $K(P|S) \leq K(x|S) + O(1) \leq K(x|S) + d \leq \log |S| - c$ , for large enough values of  $d$ . We have

$$\begin{aligned} \Pr_S(P) &= (|S| - 1)/|S| \\ &= 1 - 1/|S| \\ &= 1 - 2^{-\log |S|} \\ &\geq 1 - 2^{-c-K(P|S)}. \end{aligned}$$

Since  $S$  is  $(c, x)$ -consistent and  $P$  occurs with high probability in  $S$ , we have  $x \in P$ . But by construction,  $x \notin P$ . Therefore,  $\delta(x|S) \leq c + d$ .  $\square$

As a result,  $\text{nsoph}_c(x)$  roughly measures the minimum complexity of  $(c, x)$ -consistent sets. Also, if  $S$  is a witness to  $\text{nsoph}_c(x)$ , we can use  $S$  to infer information about  $x$ . If we want to see if  $x$  has a property  $P$ , we take many elements of  $S$  at random. If all of those elements are in  $P$ , then it is very likely that  $x$  is in  $P$  as well. Otherwise, we cannot tell whether  $x \in P$  or  $x \notin P$ . That is, witnesses to naive sophistication generally only prevent false positives, not false negatives.

## 6 Relation to Computational Depth

Computational depth measures how much harder compression is in the presence of time bounds:

$$\text{depth}^t(x) = K^t(x) - K(x)$$

where  $K^t(x) = \min_p \{ |p| : U(p, \varepsilon) = x \text{ in at most time } t \}$  is the time-bounded Kolmogorov complexity. It is well known that sophistication and computational depth are related [4,6], as are computational depth and randomness deficiency [8]. Antunes and Fortnow [6] used busy beaver computational depth:

$$\text{depth}_{\text{BB}}(x) = \min_t \{ K(t) + \text{depth}^t(x) \}.$$

They showed that coarse sophistication is equivalent to busy beaver computational depth, up to a logarithmic additive term:

$$\text{csoph}(x) \approx \text{depth}_{\text{BB}}(x).$$

We can show a similar result about naive sophistication and how it relates to a variant of busy beaver computational depth, but first we need to strengthen symmetry of information with explicit time bounds.

**Lemma 4 (Time-bounded symmetry of information).** *Symmetry of information still holds in the presence of a time bound, but the time bound increases. Let  $t$  be a time bound. Then there exists a  $t' \geq t$  such that*

$$K^{t'}(y) + K^{t'}(x|y) \leq K^t(x, y) + O(\log K^t(x, y)).$$

Furthermore,  $t' = \gamma \cdot t \log t$  where  $K(\gamma) \leq O(\log K^t(x, y))$ .

*Proof.* Let  $m = K^t(x, y)$ . Let  $V = \{ v : K^t(v, y) \leq m \}$ , and  $U = \{ u : \exists \geq |V| v. K^t(v, u) \leq m \}$ . Note that  $|U| \leq 2^m/|V|$ , that  $x \in V$  and  $y \in U$ , and that both  $V$  and  $U$  are enumerable in time  $t' = 2^{2m+O(1)} \cdot t \log t$ , and we have  $K(2^{2m+O(1)}) \leq O(\log m)$ . Finally, we have,

$$\begin{aligned} K^{t'}(y) + K^{t'}(x|y) &\leq (K^{t'}(U) + \log |U|) + (K^{t'}(V|y) + \log |V|) + O(1) \\ &\leq \log |U| + \log |V| + O(\log m) \\ &\leq (m - k) + k + O(\log m) \\ &\leq K^t(x, y) + O(\log K^t(x, y)). \end{aligned} \quad \square$$



**Theorem 5 (Naive sophistication as computational depth).** *Up to a logarithmic additive term, we have:*

$$\text{ncsoph}(x) = \min_t \{ K(t) + \text{depth}^t(x|t) \}.$$

*Proof.*

( $\leq$ ) Let  $t$  minimize the right-hand side. Let  $k = K^t(x|t)$ , and let  $S = \{ y : K^t(y|t) \leq k \}$ . We have  $\log |S| \leq k$  and  $K(S|t) \leq K(k) \leq O(\log |x|)$ . Therefore,

$$\begin{aligned} \text{ncsoph}(x) &\leq K(S) + \log |S| - K(x|S) \\ &\leq (K(t) + O(\log |x|)) + k - (K(x|t) - O(\log |x|)) \\ &\leq K(t) + K^t(x|t) - K(x|t) + O(\log |x|). \end{aligned}$$

( $\geq$ ) Let  $S$  be a witness to  $\text{ncsoph}(x)$ . Let  $t$  be a time bound sufficiently large in order to describe  $S$  fully, pick any given element from it, and also to return  $t$ , in such a way that  $K^t(x, t) \leq K(S) + \log |S| + O(1)$ . For any  $S$ , we can construct such a  $t$ , so  $K(t|S) \leq O(\log |x|)$ . Using Lemma 4, we obtain a time bound  $t'$  such that  $K(t) + K^{t'}(x|t) \leq K^t(x, t) + O(\log |x|)$ , with  $K(t'|t) = O(\log |x|)$  and  $K(t|t') = O(\log |x|)$ . We extend this with a time bound  $t'' = c \cdot t'$  for some constant  $c$ , such that  $K^{t''}(x|t'') \leq K^{t'}(x|t) + O(\log |x|)$ . As a result, we have,

$$\begin{aligned} \min_t \{ K(t) + K^t(x|t) - K(x|t) \} &\leq K(t'') + K^{t''}(x|t'') - K(x|t'') \\ &\leq K(t) + K^{t'}(x|t) - K(x|t) + O(\log |x|) \\ &\leq K^t(x, t) - K(x|t) + O(\log |x|) \\ &\leq K(S) + \log |S| - K(x|S) + O(\log |x|) \\ &= \text{ncsoph}(x) + O(\log |x|). \quad \square \end{aligned}$$

This reveals an alternative definition for busy beaver computational depth, that holds up to a logarithmic additive term:

$$\begin{aligned} \text{depth}_{\text{BB}}(x) &= \min_t \{ K(t) + \text{depth}^t(x) \} \\ &\approx \min_t \{ K(t) + \text{depth}^t(x|t) \}. \end{aligned}$$

## 7 Conclusion

We have defined naive sophistication, a variant of sophistication based on randomness deficiency, and we have seen that it is equivalent to sophistication up to low order terms. Naive sophistication gives us a new perspective on sophistication and allows us to apply it in new ways, as we have done in looking at lossy compression and computational depth through new eyes. We have seen that naive sophistication arises naturally in trying to measure how much information of a string we can throw away without losing the ability to query its properties (without false positives). We have also seen that naive sophistication allows us to find an alternative definition for busy beaver computational depth.

**Acknowledgements.** We wish to thank Péter Gács, Paul Vitányi, Nikolay Vereshchagin, and the anonymous reviewers for their insightful comments. This work was partially supported by the national science foundation of Portugal, Fundação para a Ciência e Tecnologia, through the project CSI<sup>2</sup> with the reference PTDC/EIA-CCO/099951/2008 and through grants of the Instituto de Telecomunicações. André Souto was also supported by Fundação para a Ciência e Tecnologia through the scholarship SFRH/BPD/76231/2011.

## References

1. Vereshchagin, N., Vitányi, P.: Kolmogorov's structure functions and model selection. *IEEE Transactions on Information Theory* 50(12), 3265–3290 (2004)
2. Li, M., Vitányi, P.: An introduction to Kolmogorov complexity and its applications, 3rd edn. Springer (2008)
3. Gács, P.: On the symmetry of algorithmic information. *Soviet Mathematics Doklady* 15, 1477–1480 (1974)
4. Koppel, M.: Structure. In: Herken, R. (ed.) *The Universal Turing Machine: A Half-Century Survey*, pp. 435–452. Oxford University Press (1988)
5. Vitányi, P.: Meaningful information. *IEEE Transactions on Information Theory* 52(10), 4617–4626 (2006)
6. Antunes, L., Fortnow, L.: Sophistication revisited. *Theory of Computing Systems* 45(1), 150–161 (2009)
7. Aaronson, S.: Can a string's sophistication be defined in an unsophisticated way? (2012), <http://mathoverflow.net/questions/103619>
8. Antunes, L., Matos, A., Souto, A., Vitányi, P.: Depth as randomness deficiency. *Theory of Computing Systems* 45(4), 724–739 (2009)