Computational Depth of Infinite Strings Revisited

Antunes, L.; Costa, A.; Matos, A.; Vitanyi, P.

Published in:
Computation and Logic in the Real World

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)
Computational Depth of Infinite Strings Revisited

Luís Antunes¹ *, Armindo Costa², Armando Matos¹, and Paul Vitányi³

¹ DCC-UP & LIACC
² Durham University
³ CWI & University of Amsterdam

Abstract. Bennett introduced the notion of logical depth of an object as the amount of time required for an algorithm to derive the object from a shorter description. He also defined logical depth for infinite strings, in particular strongly and weakly deep sequences. Later Lutz, Juedes and Mayordomo have further studied and related these measures. Recently Antunes et al. noted that logical depth, as introduced by Bennett is connected to Kolmogorov and Levin’s notion of “randomness deficiency”. Based on this connection, we revisit the notion of computational depth for infinite strings, introducing the notion of super deep sequences and relate it with other approaches.

1 Introduction

The Kolmogorov complexity of a string is a rigorous measure of the amount of information contained in it. A string with high Kolmogorov complexity contains lots of information. A randomly generated string has, with high probability, high Kolmogorov complexity and hence is very informative. However, intuitively, the very fact that it is random makes it unlikely to have useful and meaningful information. How can we measure the nonrandom information in a string?

Bennett [Ben88] introduced the notion of logical depth of an object as the amount of time required for an algorithm to derive the object from a shorter description. In fact, with some probability, we can derive the object by simply flipping a coin. But for long objects this probability is small. If the object has a short description then we can obtain it by flipping a fair coin with higher probability. In order to solve some stability problems, Bennett’s definition considers not only the shortest description of the object, but all descriptions of it that have nearly minimal length.

Later Antunes et al. [AFMV06] defined a simpler notion by taking the difference between polynomial-time Kolmogorov complexity and traditional unbounded Kolmogorov complexity. We have seen a number of results about computational depth such as: the use of depth to characterize the worst-case running time of problems that run quickly on average over all polynomial-time samplable distributions [AF05], give a generalization of sparse and random sets [AFMV06], as well as find satisfying assignments for formulas that have at least one assignment of logarithmic depth [AFPS06].

For infinite sequences Bennett identified the classes of weekly and strongly deep sequences, and showed that the halting problem is strongly deep. Subsequently Judes, Lathrop, and Lutz [JLL94] extended Bennett’s work defining the classes of weekly useful sequences and proved that every weakly useful sequence is strongly deep in the sense of Bennett. Fenner et al. [FLMR05] proved that there exist sequences that are weakly useful but not strongly useful. Lathrop and Lutz [LL99] introduced refinements (named recursive weak depth and recursive strong depth) of Bennett’s notion of weak and strong depth, and studied its fundamental properties, showing that recursively weakly (strongly) deep sequences form a proper subclass of

* Partially supported by KCrypt (POSC/EIA/60819/2004) and funds granted to LIACC through the Programa de Financiamento Pluriannual, FCT and Programa POSI.
the class of weakly (strongly) deep sequences, and also that every weakly useful sequences is recursive strongly deep.

Antunes et al. [ACMV07] noted that logical depth and computational depth are instances of a more general measure namely, the randomness deficiency of a string $x$ with respect to a probability distribution, developed by Levin [Lev84]. Based on this connection we extend the notion of computational depth to infinite strings, introducing the notion of super deep sequences and relate it with other approaches. The paper is organized as follows: in Section 2 we introduce the necessary notation and definitions, in Section 3 we study the common information for infinite strings and finally in Section 4, based on the connection between randomness deficiency and computational depth, we introduce and study the notion of super deep for infinite sequences.

2 Preliminaries

We briefly introduce Kolmogorov complexity. We refer to the textbook by Li and Vitányi [LV97] for more details.

**Definition 1.** Let $U$ be a fixed prefix free universal Turing machine; for every string $x \in \{0,1\}^*$, the Kolmogorov complexity of $x$ is, $K(x) = \min \{|p| : U(p) = x\}$.

For any time constructible $t$, the $t$-time-bounded Kolmogorov complexity of $x$ is, $K^t(x) = \min \{|p| : U(p) = x \text{ in at most } t(|x|) \text{ steps}\}$.

A different universal machine $U$ may affect the program size $|p|$ by at most an additive constant factor, and the running time $t$ by at most a logarithmic multiplicative factor. The same holds for all other measures we will introduce.

We refer to mutual information of two finite strings as

$$I(x : y) = K(x) + K(y) - K(x, y)$$

and to algorithmic information that a finite string $x$ as about another finite string $y$ as

$$I(x ; y) = K(x) - K(x|y).$$

**Depth of finite strings** After some attempts Bennett [Ben88] formally defined the $b$-significant logical depth of an object $x$ as the time required by a standard universal Turing machine to generate $x$ by a program that is no more than $b$ bits longer than the shortest descriptions of $x$. Formally:

**Definition 2 (Bennett).** The logical depth of a string $x$ at a significance level $b$ is

$$ldepth_b(x) = \min \left\{ t : \frac{Q^t_U(x)}{Q_U(x)} \geq 2^{-b} \right\}$$

where $U^t(p) = x$ means that $U$ computes $x$ within $t$ steps and halts and $Q^t_U(x) = \sum_{U^t(p) = x} 2^{-|p|}$.

Antunes et al. [AFMV06] developed the notion of Computational Depth in order to capture the amount of non-random or useful information in a string. The concept is simple: they consider the difference of two different Kolmogorov complexity measures. What remains is the "nonrandom" or "useful" information.

**Definition 3 (Computational Depth).** Let $t$ be a constructible time bound. For any string $x \in \{0,1\}^*$,

$$\text{depth}_t(x) = K^t(x) - K(x).$$
It is interesting to note that logical depth and computational depth are all instances of a more general measure, namely the randomness deficiency of a string \( x \) with respect to a probability distribution as developed by Levin [Lev84].

**Definition 4 (Levin).** For any r.e. measure \( \mu \),

\[
d(x/\mu) = \left\lfloor \log \frac{m(x)}{\mu(x)} \right\rfloor
\]

is the randomness deficiency\(^4\) of \( x \) with respect to \( \mu \), where \( m(x) = 2^{-K(x)} \) is the universal distribution.

Levin showed that the randomness deficiency of \( x \) with respect to \( \mu \) is the largest randomness \( \mu \)-test for \( x \). So \( d(x/\mu) \) is, in a sense, a universal characterization of “non-random” or “useful” information in a string \( x \) with respect to the measure \( \mu \).

**Depth of infinite strings** We denote the \( n \)-prefix of an infinite string \( \alpha \) by \( \alpha_n \), the \( i \)-bit of the string \( x \) by \( x_i \). Bennett also introduced the classes of weakly and strongly deep sequences and proved that the halting problem is strongly deep.

**Definition 5 ([Ben88]).** An infinite binary sequence \( \alpha \) is defined as

- weakly deep if it is not computable in recursively bounded time from any algorithmically random infinite sequence.
- strongly deep if at every significance level \( b \), and for every recursive function \( t \), all but finitely many initial segments \( \alpha_n \) have depth exceeding \( t(n) \).

The relation between depth and usefulness was studied by Juedes, Lathrop and Lutz [JLL94] who defined the conditions for weak and strong usefulness and showed that every weakly useful sequence is strongly deep. This result generalizes Bennett’s remark on the depth of diagonal of the halting problem, strengthening the relation between depth and usefulness. Latter Fenner et al. [FLMR05] proved the existence of sequences that are weakly useful but not strongly useful.

**Definition 6.** An infinite binary sequence \( \alpha \) is defined as

- strongly useful if there is a computable time bound within which every decidable sequence is Turing reducible to \( \alpha \); and
- weakly useful if there is a computable time bound within which all the sequences in a non-measure 0 subset of the set of decidable sequences are Turing reducible to \( \alpha \).

Levin’s notion of randomness deficiency can be extended to infinite strings. Let \( \alpha \) and \( \beta \) be two sequences and \( m \otimes m \) be defined by \( m \otimes m(\alpha, \beta) = m(\alpha)m(\beta) \).

**Definition 7 (Levin).** The value \( D(\alpha/\mu) = \lfloor \log \sup(m(\alpha_n)/\mu(\alpha_n)) \rfloor \) is called the randomness deficiency of \( \alpha \) with respect to the semi-measure \( \mu \).

**Definition 8 (Levin).** The value \( I(\alpha : \beta) = D((\alpha, \beta)/m \otimes m) \) is called the amount of information in \( \alpha \) about \( \beta \) or the deficiency of their independence.

This definition is equivalent to the mutual information \( I(\alpha : \beta) = \sup I(\alpha_n : \beta_n) \).

\(^4\) \( [r] \) denotes the integer part of \( r \) and \( \lceil \alpha \rceil \) denotes the smallest integer bigger than \( \alpha \).
3 On the Information of Infinite Strings

Let $\alpha$ and $\gamma$ be two random infinite and independent strings (in the sense that their prefixes are independent). Consider the following sequence $\beta$

$$
\beta = \alpha^1 \gamma^1 \alpha^2 \gamma^2 \ldots
$$

By Definition 8 we have

$$
I(\alpha : \beta) = \sup I(\alpha_n : \beta_n)
= \sup(K(\alpha_n) + K(\beta_n) - K(\alpha_n, \beta_n))
\geq \sup(K(\alpha_n/2)) = \infty.
$$

As $I(\beta : \alpha) = I(\alpha : \beta)$ then $I(\beta : \alpha) = \infty$. However, intuitively $\beta$ contains more information about $\alpha$ than the other way around. If one considers both sequences as the limit of increasing prefixes that are growing at the same speed, then, their mutual information should be half of the full information, that is, the information of an incompressible sequence.

This seems to be a lacuna in Definition 8. It says more when the information is finite but that is precisely when we don’t need an accurate result, if it is finite we can argue that they are independent. One should be able to classify the cases where the mutual information is infinite. Two infinite strings may have infinite mutual information and yet infinite information still lacks in order to reconstruct each of them. In the previous example $\alpha$ fails to provide all the information of $\gamma$, which has infinite information. In this section we will present three approaches to do it. In order to have a proportion of information as the prefixes grow up we need some normalization in the process.

3.1 The Algorithmic Information Point of View

Assumed that when calculating the information that a string $\alpha$ has about a string $\beta$ one have the full knowledge of the first. We are looking for a normalized algorithmic information measure $In$ that applied the Example 1 gives

$$
In(\alpha; \alpha) = 1; \quad In(\alpha; \beta) = 1/2; \quad In(\beta; \alpha) = 1; \quad In(\beta; \beta) = 1
$$

Definition 9 (First attempt). Given two infinite strings $\alpha$ and $\beta$ the normalized algorithmic information that $\beta$ has about $\alpha$ is defined as

$$
In(\beta; \alpha) = \lim_{n \to \infty} \lim_{m \to \infty} \frac{I(\beta_m; \alpha_n)}{I(\alpha_n; \alpha_n)}
$$

One drawback of this definition is that the limit does not always exist. However, it does exist for the Example 1 with the desired properties. Furthermore, supposing that the above limits exist we obtain for the same $\alpha$ and $\beta$

$$
In(\alpha; \alpha) = 1; \quad In(\beta; \beta) = 1
$$

$$
In(\alpha; \beta) = \lim_{n \to \infty} \frac{K(\alpha_{n/2})}{K(\alpha_{n/2}) + K(\gamma_{n/2})} = \lim_{n \to \infty} \frac{K(\alpha_n)}{K(\alpha_n) + K(\gamma_n)}; \quad In(\beta; \alpha) = 1
$$

Showing that the proportion of the information that $\alpha$ gives on $\beta$ ought to depend on the complexity of $\gamma$. It seems clear that when $I(x : y) = \infty$ then $In(x : y) \neq 0$. Thus one continue to obtain a characterization of independent sequences ($I(x : y) < \infty$) and already can say new things when $I(x : y) = \infty$ and the above limits exist.
Definition 10. Given two sequences $\alpha$ and $\beta$ we define the lower normalized algorithmic information that $\beta$ has about $\alpha$ as

$$I_*(\beta; \alpha) = \liminf_{n \to \infty} \lim_{m \to \infty} \frac{I(\beta_m; \alpha_n)}{I(\alpha_n; \alpha_n)}$$

and the upper normalized algorithmic information that $\beta$ has about $\alpha$ as

$$I^*(\beta; \alpha) = \limsup_{n \to \infty} \lim_{m \to \infty} \frac{I(\beta_m; \alpha_n)}{I(\alpha_n; \alpha_n)}$$

3.2 The mutual information point of view

In this subsection we define the common information between two sequences based on the mutual information, looking to both prefixes growing at the same rate.

Definition 11. Given two sequences $\alpha$ and $\beta$ we define the lower normalized mutual information that $\beta$ has about $\alpha$ as

$$I_*(\beta : \alpha) = \liminf_{n \to \infty} \frac{I(\beta_n : \alpha_n)}{I(\alpha_n : \alpha_n)}$$

and the upper normalized mutual information that $\beta$ has about $\alpha$ as

$$I^*(\beta : \alpha) = \limsup_{n \to \infty} \frac{I(\beta_n : \alpha_n)}{I(\alpha_n : \alpha_n)}$$

We now define independent sequences.

Definition 12. Two infinite strings, $\alpha$ and $\beta$, are independent if $I^*(\alpha : \beta) = I^*(\beta : \alpha) = 0$.

In [Lut00, Lut02], the author developed a constructive version of Hausdorff dimension. That dimension assigns to every binary sequence $\alpha$ a real number $\dim(\alpha)$ in the interval $[0,1]$. Lutz claims that the dimension of a sequence is a measure of its information density, the idea is to differentiate sequences by non-randomness degrees, namely by their dimension. Our approach is precisely to introduce a measure of the density of information that one sequence has about the other, in the total amount of the other's information. So we differentiate non-independent sequences, by their normalized mutual information. Considering Example 1 we have,

$$I_n(\alpha : \beta) = 1; \quad I_n(\beta : \alpha) = 1$$

$$I_n(\alpha : \beta) = \lim_{n \to \infty} \frac{K(\alpha_{n/2})}{K(\alpha_n)} = \frac{1}{2}$$

$$I_n(\beta : \alpha) = \lim_{n \to \infty} \frac{K(\beta_{n/2})}{K(\beta_n)} = \lim_{n \to \infty} \frac{K(\alpha_{n/2})}{K(\alpha_n/2) + K(\gamma_{n/2})} = \frac{1}{2}$$

Mayordomo [May02] redefined constructive Hausdorff dimension in terms of Kolmogorov complexity.

Theorem 1 (Mayordomo). For every sequence $\alpha$,

$$\dim(\alpha) = \liminf_{n \to \infty} \frac{K(\alpha_n)}{n}$$
So, now the connection between constructive dimension and normalized information measure introduced here is clear. It is only natural to accomplish results about the Hausdorff constructive dimension of a sequence, knowing the dimension of another, and their normalized information.

**Lemma 1.** Let \( \alpha \) and \( \beta \) be two sequences. Then

\[
I_*(\alpha : \beta), \dim(\beta) \leq \dim(\alpha)
\]

**Proof.** Obviously \( 0 \leq I_*(\alpha : \beta) \leq 1 \) and \( 0 \leq \dim(\beta) \leq 1 \). So,

\[
I_*(\alpha : \beta), \dim(\beta) = \left( \lim_{n \to \infty} \frac{I(\alpha_n : \beta_n)}{I(\beta_n : \beta_n)} \right) \left( \lim_{n \to \infty} \frac{K(\beta_n)}{n} \right)
\]

\[
\leq \lim_{n \to \infty} \frac{I(\alpha_n : \beta_n) K(\beta_n)}{n}
\]

\[
= \lim_{n \to \infty} \frac{I(\alpha_n : \beta_n)}{n}
\]

\[
\leq \lim_{n \to \infty} \frac{K(\alpha_n)}{n}
\]

\[
= \dim(\alpha)
\]

### 3.3 The Hausdorff constructive dimension point of view

In this subsection we define the common information between two sequences based on Hausdorff constructive dimension.

**Definition 13.** The dimensional mutual information of the sequences \( \alpha \) and \( \beta \) is defined as

\[
I_{\dim}(\alpha, \beta) = \dim(\alpha) + \dim(\beta) - 2 \dim(\langle \alpha, \beta \rangle)
\]

This information measure is symmetric, the definitions considers twice \( \dim(\langle \alpha, \beta \rangle) \) because when encoding the prefixes \( \alpha_n \) and \( \beta_n \) the result is a 2n-length string.

As, \( I_{\dim}(\alpha, \beta) = \dim(\alpha) + \dim(\beta) - 2 \dim(\langle \alpha, \beta \rangle) \)

\[
= \lim_{n \to \infty} \frac{K(\alpha_{n/2})}{n/2} + \lim_{n \to \infty} \frac{K(\beta_{n/2})}{n/2} - 2 \lim_{n \to \infty} \frac{K(\langle \alpha, \beta \rangle_n)}{n}
\]

\[
\leq \lim_{n \to \infty} \frac{K(\alpha_{n/2}) + K(\beta_{n/2}) - K(\langle \alpha, \beta \rangle_{n/2})}{n/2}
\]

\[
= \lim_{n \to \infty} \frac{I(\alpha_n : \beta_n)}{n}
\]

\[
\leq \min(\dim(\alpha), \dim(\beta))
\]

then the following lemma holds.

**Lemma 2.** Let \( \alpha \) and \( \beta \) be two sequences. Then

\[
I_{\dim}(\alpha, \beta) \leq \min(\dim(\alpha), \dim(\beta))
\]

When the limits in Hausdorff constructive dimension exist (in the sense that we take \( \lim_{n \to \infty} \) instead \( \lim_{n \to \infty} \)) we have

\[
I_{\dim}(\alpha, \beta) = I_*(\alpha : \beta), \dim(\beta) = I_*(\beta : \alpha), \dim(\alpha)
\]

One can easily modify the definitions introduced in this section by considering the limits when \( n \) goes to the length of the string, or the maximum length of the strings being considered. One should also notice that when \( x \) and \( y \) are finite strings and \( K(y) \geq K(x) \), \( I_*(x : y) \) is \( 1 - d(x, y) \), where \( d(x, y) \) is the normalized information distance studied in [Li03].
4 Depth on infinite strings

The Hausdorff constructive dimension comes hand to hand with the three information theories for infinite strings studied before. Therefore, in this section we define the dimensional computational depth of a sequence in order to study the nonrandom information on an infinite string.

Definition 14. The dimensional depth of a sequence $\alpha$ is defined as

$$\text{depth}^t_{\text{dim}}(\alpha) = \liminf_{n \to \infty} \frac{D(\alpha_n/Q^t(\alpha_n))}{n}.$$ 

Definition 15. The $t$-bounded dimension of a sequence $\alpha$ is defined as

$$\text{dim}^t(\alpha) = \liminf_{t \to \infty} -\frac{\log Q^t(\alpha_n)}{n}.$$ 

Lemma 3.

$$\text{depth}^t_{\text{dim}}(\alpha) \leq \text{dim}^t(\alpha) - \text{dim}(\alpha)$$

Proof.

$$\text{depth}^t_{\text{dim}}(\alpha) = \liminf_{n \to \infty} \frac{D(\alpha_n/Q^t(\alpha_n))}{n}$$

$$= \liminf_{n \to \infty} -\frac{\log Q^t(\alpha_n) - K(\alpha_n)}{n}$$

$$\leq \text{dim}^t(\alpha) - \text{dim}(\alpha).$$

Now we replace the fixed significance level $s$ by a significance function, $s : \mathbb{N} \to \mathbb{N}$ in the definition of strongly deep sequences. Naturally, we want $s(n)$ to grow very slowly so assume that $s = o(n)$. Replacing the fixed significance level $s$ in the definition of strongly deep by this significance function we obtain a tighter definition as deepness decreases with the increase of the significance level.

Definition 16. A sequence is called super deep if for every significance function $s : \mathbb{N} \to \mathbb{N}$, such that $s = o(n)$, and for every recursive function $t : \mathbb{N} \to \mathbb{N}$, all but finitely many initial segments $\alpha_n$ have depth exceeding $t(n)$.

We can in fact characterize super deep sequences using their dimensional depth. For that we need the following result in [ACMV07] stating that for any constructible time bound $t$

$$l\text{depth}_t(x) = t \Leftrightarrow \text{depth}^t(x) \leq b \text{ and } \text{depth}^{t-1}(x) > b.$$ 

Theorem 2. A sequence $\alpha$ is super deep if and only if $\text{depth}^t_{\text{dim}}(\alpha) > 0$ for all recursive time bound $t$.

Proof. Let $\alpha$ be a super deep sequence. Then for every significance function $s$, such that $s = o(n)$ and every recursive function $t$ we have that for almost all $n$, $l\text{depth}_{s(n)}(\alpha_n) > t(n)$. Then

$$\text{depth}^{t(n)}(\alpha_n) > s(n).$$

Now if for some time bound $g$, $\text{depth}^g_{\text{dim}}(\alpha) = 0$ then there exists a bound $S$, such that $S = o(n)$, and infinitely often

$$\text{depth}^{t(n)}(\alpha_n) < S(n).$$

This is absurd and therefore for all recursive time bound $t$, $\text{depth}^t_{\text{dim}}(\alpha) > 0$. 

Conversely if $\text{depth}_{\dim}(\alpha) > 0$ then there is some $\epsilon > 0$ such that for almost all $n$, $\text{depth}_{\dim}(\alpha_n) > \epsilon n$. This implies that

$$l\text{depth}_{s(n)}(\alpha_n) > l\text{depth}_{\epsilon n}(\alpha_n) > t(n)$$

for all significance function $s = o(n)$ and almost all $n$. So $\alpha$ is super deep.

In [JLL94] several characterizations of strong computational depth are obtained. Following the ideas in [JLL94], we can prove analogous characterizations for super deepness.

**Theorem 3.** For every sequence $\alpha$ the following conditions are equivalent.

1. $\alpha$ is super deep.
2. For every recursive time bound $t : \mathbb{N} \to \mathbb{N}$ and every significance function $g = o(n)$, $\text{depth}_t(\alpha_n) > g(n)$ a.e.
3. For every recursive time bound $t : \mathbb{N} \to \mathbb{N}$ and every significance function $g = o(n)$, $K_t(\alpha_n) - K(\alpha_n) > g(n)$ a.e.
4. For every recursive time bound $t : \mathbb{N} \to \mathbb{N}$ and every significance function $g = o(n)$, $Q(\alpha_n) \geq 2^{g(n)}Q_t(\alpha_n)$ a.e.

In [JLL94] the authors proved that every weakly useful sequence is strongly deep. Following the ideas in [JLL94] we can also prove that that every weakly useful sequence is super deep.

**Theorem 4.** Every weakly useful sequence is super deep.

**Corollary 1.** The characteristic sequences of the halting problem and the diagonal halting problem are super deep.

**References**


