

# Time-Bounded Kolmogorov Complexity and Solovay Functions

Rupert Hölzl, Thorsten Kräling, and Wolfgang Merkle

Institut für Informatik, Ruprecht-Karls-Universität,  
Heidelberg, Germany

**Abstract.** A Solovay function is a computable upper bound  $g$  for prefix-free Kolmogorov complexity  $K$  that is nontrivial in the sense that  $g$  agrees with  $K$ , up to some additive constant, on infinitely many places  $n$ . We obtain natural examples of Solovay functions by showing that for some constant  $c_0$  and all computable functions  $t$  such that  $c_0 n \leq t(n)$ , the time-bounded version  $K^t$  of  $K$  is a Solovay function.

By unifying results of Bienvenu and Downey and of Miller, we show that a right-computable upper bound  $g$  of  $K$  is a Solovay function if and only if  $\Omega_g$  is Martin-Löf random. Letting  $\Omega_g = \sum 2^{-g(n)}$ , we obtain as a corollary that the Martin-Löf randomness of the various variants of Chaitin's  $\Omega$  extends to the time-bounded case in so far as  $\Omega_{K^t}$  is Martin-Löf random for any  $t$  as above.

As a step in the direction of a characterization of  $K$ -triviality in terms of jump-traceability, we demonstrate that a set  $A$  is  $K$ -trivial if and only if  $A$  is  $O(g(n) - K(n))$ -jump traceable for all Solovay functions  $g$ , where the equivalence remains true when we restrict attention to functions  $g$  of the form  $K^t$ , either for a single or all functions  $t$  as above.

Finally, we investigate the plain Kolmogorov complexity  $C$  and its time-bounded variant  $C^t$  of initial segments of computably enumerable sets. Our main theorem here is a dichotomy similar to Kummer's gap theorem and asserts that every high c.e. Turing degree contains a c.e. set  $B$  such that for any computable function  $t$  there is a constant  $c_t > 0$  such that for all  $m$  it holds that  $C^t(B \upharpoonright m) \geq c_t \cdot m$ , whereas for any nonhigh c.e. set  $A$  there is a computable time bound  $t$  and a constant  $c$  such that for infinitely many  $m$  it holds that  $C^t(A \upharpoonright m) \leq \log m + c$ . By similar methods it can be shown that any high degree contains a set  $B$  such that  $C^t(B \upharpoonright m) \geq^+ m/4$ . The constructed sets  $B$  have low unbounded but high time-bounded Kolmogorov complexity, and accordingly we obtain an alternative proof of the result due to Juedes, Lathrop, and Lutz [JLL] that every high degree contains a strongly deep set.

## 1 Introduction and Overview

Prefix-free Kolmogorov complexity  $K$  is not computable and in fact does not even allow for computable lower bounds. However, there are computable upper bounds for  $K$  and, by a construction that goes back to Solovay [BD, S], there are even computable upper bounds that are nontrivial in the sense that  $g$  agrees

with  $K$ , up to some additive constant, on infinitely many places  $n$ ; such upper bounds are called Solovay functions.

For any computable time-bound  $t$ , the time-bounded version  $K^t$  of  $K$  is obviously a computable upper bound for  $K$ , and we show that  $K^t$  is indeed a Solovay function in case  $c_0 n \leq t(n)$  for some appropriate constant  $c_0$ . As a corollary, we obtain that the Martin-Löf randomness of the various variants of Chaitin's  $\Omega$  extends to the time-bounded case in so far as for any  $t$  as above, the real number

$$\Omega_{K^t} = \sum_{n \in \mathbb{N}} \frac{1}{2^{K^t(n)}}$$

is Martin-Löf random. The corresponding proof exploits the result by Bienvenu and Downey [BD] that a computable function  $g$  such that  $\Omega_g = \sum 2^{-g(n)}$  converges is a Solovay function if and only if  $\Omega_g$  is Martin-Löf random. In fact, this equivalence extends by an even simpler proof to the case of functions  $g$  that are just right-computable, i.e., effectively approximable from above, and one then obtains as special cases the result of Bienvenu and Downey and a related result of Miller where the role of  $g$  is played by the fixed right-computable but noncomputable function  $K$ .

An open problem that received some attention recently [BDG, DH, N] is whether the class of  $K$ -trivial sets coincides with the class of sets that are  $g(n)$ -jump-traceable for all computable functions  $g$  such that  $\sum 2^{-g(n)}$  converges. As a step in the direction of a characterization of  $K$ -triviality in terms of jump-traceability, we demonstrate that a set  $A$  is  $K$ -trivial if and only if  $A$  is  $O(g(n) - K(n))$ -jump traceable for all Solovay functions  $g$ , where the equivalence remains true when we restrict attention to functions  $g$  of the form  $K^t$ , either for a single or all functions  $t$  as above.

Finally, we consider the time-bounded and unbounded Kolmogorov complexity of the initial segments of sets that are computationally enumerable, or c.e., for short. The initial segments of a c.e. set  $A$  have small Kolmogorov complexity, more precisely, by Barzdins' lemma it holds that  $C(A \upharpoonright m) \leq^+ 2 \log m$ , where  $C$  denotes plain Kolmogorov complexity. Theorem 4, our main result in this section, has a structure similar to Kummer's gap theorem in so far as it asserts a dichotomy in the complexity of initial segments between high and nonhigh c.e. sets. More precisely, every high c.e. Turing degree contains a c.e. set  $B$  such that for any computable function  $t$  there is a constant  $c_t > 0$  such that for all  $m$  it holds that  $C^t(B \upharpoonright m) \geq c_t \cdot m$ , whereas for any nonhigh c.e. set  $A$  there is a computable time bound  $t$  and a constant  $c$  such that for infinitely many  $m$  it holds that  $C^t(A \upharpoonright m) \leq \log m + c$ . By similar methods it can be shown that any high degree contains a set  $B$  such that  $C^t(B \upharpoonright m) \geq^+ m/4$ . The constructed sets  $B$  have low unbounded but high time-bounded Kolmogorov complexity, and accordingly we obtain an alternative proof of the result due to Juedes, Lathrop, and Lutz [JLL] that every high degree contains a strongly deep set.

*Notation.* In order to define plain and prefix-free Kolmogorov complexity, we fix additively optimal oracle Turing machines  $\mathbb{V}$  and  $\mathbb{U}$ , where  $\mathbb{U}$  has prefix-free domain. We let  $C^A(x)$  denote the Kolmogorov-complexity of  $x$  with respect

to  $\mathbb{V}$  relative to oracle  $A$ , let  $C(x) = C^{\emptyset}(x)$ , and similarly define prefix-free Kolmogorov complexity  $K^A$  and  $K$  with respect to  $\mathbb{U}$ . In connection with the definition of time-bounded Kolmogorov complexity, we assume that  $\mathbb{V}$  and  $\mathbb{U}$  both are able to simulate any other Turing machine  $M$  running for  $t$  steps in  $O(t \cdot \log(t))$  steps for an arbitrary machine  $M$  and in  $O(t(n))$  steps in case  $M$  has only two work tapes.

For a computable function  $t : \mathbb{N} \rightarrow \mathbb{N}$  and a machine  $M$ , the Kolmogorov complexity relative to  $M$  with time bound  $t$  is

$$C_M^t(n) := \min\{|\sigma| : M(\sigma) \downarrow = n \text{ in at most } t(|n|) \text{ steps}\},$$

and we write  $C^t$  for  $C_{\mathbb{U}}^t$ . The prefix-free Kolmogorov complexity with time bound  $t$  denoted by  $K_M^t(n)$  and  $K^t(n) = K_{\mathbb{U}}^t(n)$  is defined likewise by considering only prefix-free machines and the corresponding universal machine  $\mathbb{U}$  in place of  $\mathbb{U}$ .

We identify strings with natural numbers by the order isomorphism between the length-lexicographical order on strings and the usual order on  $\mathbb{N}$ , and we write  $|m|$  for the length of the string that corresponds to the number  $m$ , where then  $|m|$  is roughly  $\log m$ .

## 2 Solovay Functions and Martin-Löf Randomness

**Definition 1 (Li, Vitányi [LV]).** *A computable function  $f : \mathbb{N} \rightarrow \mathbb{N}$  is called a Solovay function if  $K(n) \leq^+ f(n)$  for all  $n$  and  $K(n) =^+ f(n)$  for infinitely many  $n$ .*

Solovay [S, BD] had already constructed Solovay functions and by slightly varying the standard construction, next we observe that time-bounded prefix-free Kolmogorov complexity indeed provides natural examples of Solovay functions.

**Theorem 1.** *There is a constant  $c_0$  such that time-bounded prefix-free Kolmogorov complexity  $K^t$  is a Solovay function for any computable function  $t : \mathbb{N} \rightarrow \mathbb{N}$  such that  $c_0 n \leq t(n)$  holds for almost all  $n$ .*

*Proof.* Fix a standard effective and effectively invertible pairing function  $\langle \cdot, \cdot \rangle : \mathbb{N}^2 \rightarrow \mathbb{N}$  and define a tripling function  $[\cdot, \cdot, \cdot] : \mathbb{N}^3 \rightarrow \mathbb{N}$  by letting

$$[s, \sigma, n] = 1^s 0 \langle \sigma, n \rangle.$$

Let  $M$  be a Turing machine with two tapes that on input  $\sigma$  uses its first tape to simulate the universal machine  $\mathbb{U}$  on input  $\sigma$  and, in case  $\mathbb{U}(\sigma) = n$ , to compute  $\langle \sigma, n \rangle$ , while maintaining on the second tape a unary counter for the number of steps of  $M$  required for these computations. In case eventually  $\langle \sigma, n \rangle$  had been computed with final counter value  $s$ , the output of  $M$  is  $z = [s, \sigma, n]$ , where by construction in this case the total running time of  $M$  is in  $O(s)$ .

Call  $z$  of the form  $[s, \sigma, n]$  a Solovay triple in case  $M(\sigma) = z$  and  $\sigma$  is an optimal code for  $n$ , i.e.,  $K(n) = |\sigma|$ . For some appropriate constant  $c_0$  and any

computable function  $t$  that eventually is at least  $c_0n$ , for almost all such triples  $z$  it then holds that

$$K(z) =^+ K^t(z),$$

because given a code for  $M$  and  $\sigma$ , by assumption the universal machine  $\mathbb{U}$  can simulate the computation of the two-tape machine  $M$  with input  $\sigma$  with linear overhead, hence  $\mathbb{U}$  uses time  $O(s)$  plus the constant time required for decoding  $M$ , i.e., time at most  $c_0|z|$ .  $\square$

Next we derive a unified form of a characterization of Solovay function in terms of Martin-Löf randomness of the corresponding  $\Omega$ -number due to Bienvenu and Downey [BD] and a result of Miller [M] that asserts that the notions of weakly low and low for  $\Omega$  coincide. Before, we review some standard notation and facts relating to  $\Omega$ -numbers.

**Definition 2.** For a function  $f: \mathbb{N} \rightarrow \mathbb{N}$ , the  $\Omega$ -number of  $f$  is

$$\Omega_f := \sum_{n \in \mathbb{N}} 2^{-f(n)}$$

**Definition 3.** A function  $f: \mathbb{N} \rightarrow \mathbb{N}$  is an information content measure relative to a set  $A$  in case  $f$  is right-computable with access to the oracle  $A$  and  $\Omega_f$  converges; furthermore, the function  $f$  is an information content measure if it is an information content measure relative to the empty set.

The following remark describes for a given information content measure  $f$  an approximation from below to  $\Omega_f$  that has certain special properties. For the sake of simplicity, in the remark only the oracle-free case is considered and the virtually identical considerations for the general case are omitted.

*Remark 1.* For a given information content measure  $f$ , we fix as follows a non-decreasing computable sequence  $a_0, a_1, \dots$  that converges to  $\Omega_f$  and call this sequence the canonical approximation of  $\Omega_f$ .

First, we fix some standard approximation to the given information content measure  $f$  from above, i.e., a computable function  $(n, s) \mapsto f_s(n)$  such that for all  $n$  the sequence  $f_0(n), f_1(n), \dots$  is a nonascending sequence of natural numbers that converges to  $f(n)$ , where we assume in addition that  $f_s(n) - f_{s+1}(n) \in \{0, 1\}$ . Then in order to obtain the  $a_i$ , let  $a_0 = 0$  and given  $a_i$ , define  $a_{i+1}$  by searching for the next pair of the form  $(n, 0)$  or the form  $(n, s + 1)$  where in addition it holds that  $f_s(n) - f_{s+1}(n) = 1$  (with some ordering of pairs understood), let

$$d_i = 2^{-f_0(n)} \quad \text{or} \quad d_i = 2^{-f_{s+1}(n)} - 2^{-f_s(n)} = 2^{-f_s(n)},$$

respectively, and let  $a_{i+1} = a_i + d_i$ . Furthermore, in this situation, say that the increase of  $d_i$  from  $a_i$  to  $a_{i+1}$  occurs due to  $n$ .

It is well-known [DH] that among all right-computable functions exactly the information content measures are, up to an additive constant, upper bounds for the prefix-free Kolmogorov complexity  $K$ .

Theorem 2 unifies two results by Bienvenu and Downey [BD] and by Miller [M], which are stated below as Corollaries 1 and 2. The proof of the backward direction of the equivalence stated in Theorem 2 is somewhat more direct and uses different methods when compared to the proof of Bienvenu and Downey, and is quite a bit shorter than Miller’s proof, though the main trick of delaying the enumeration via the notion of a matched increase is already implicit there [DH, M]. Note in this connection that Bienvenu has independently shown that Miller’s result can be obtained as a corollary to the result of Bienvenu and Downey [DH].

**Theorem 2.** *Let  $f$  be an information content measure relative to a set  $A$ . Then  $f$  has the Solovay property with respect to  $K^A$ , i.e.,*

$$\lim_{n \rightarrow \infty} (f(n) - K^A(n)) \neq +\infty \tag{1}$$

*if and only if  $\Omega_f$  is Martin-Löf random relative to  $A$ .*

*Proof.* We first show the backwards direction of the equivalence asserted in the theorem, where the construction and its verification bear some similarities to Kučera and Slaman’s [KS] proof that left-computable sets that are not Solovay complete cannot be Martin-Löf random. We assume that (1) is false and construct a sequence  $U_0, U_1, \dots$  of sets that is a Martin-Löf test relative to  $A$  and covers  $\Omega_f$ . In order to obtain the component  $U_c$ , let  $a_0, a_1, \dots$  be the canonical approximation to  $\Omega_f$  where in particular  $a_{i+1} = a_i + d_i$  for increases  $d_i$  that occur due to some  $n$ . Let  $b_i$  be the sum of all increases  $d_j$  such that  $j \leq i$  and where  $d_j$  and  $d_i$  are due to the same  $n$ . With the index  $c$  understood, say an increase  $d_i$  due to  $n$  is matched if it holds that

$$2^{c+1}b_i \leq 2^{-K^A(n)}.$$

For every  $d_i$  for which it could be verified that  $d_i$  is matched, add an interval of size  $2d_i$  to  $U_c$  where this interval either starts at  $a_i$  or at the maximum place that is already covered by  $U_c$ , whichever is larger. By construction the sum of all matched  $d_i$  is at most  $\Omega^A/2^{c+1} \leq 2^{-(c+1)}$  and the sets  $U_c$  are uniformly c.e. relative to  $A$ , hence  $U_0, U_1, \dots$  is a Martin-Löf test relative to  $A$ . Furthermore, this test covers  $\Omega_f$  because by the assumption that (1) is false, for any  $c$  almost all increases are matched.

For ease of reference, we review the proof of the forward direction of the equivalence asserted in the theorem, which follows by the same line of standard argument that has already been used by Bienvenu and Downey and by Miller. For a proof by contraposition, assume that  $\Omega_f$  is not Martin-Löf random relative to  $A$ , i.e., for every constant  $c$  there is a prefix  $\sigma_c$  of  $\Omega_f$  such that  $K^A(\sigma_c) \leq |\sigma_c| - 2c$ . Again consider the canonical approximation  $a_0, a_1, \dots$  to  $\Omega_f$  where  $f(n, 0), f(n, 1), \dots$  is the corresponding effective approximation from above to  $f(n)$  as in Remark 1. Moreover, for  $\sigma_c$  as above we let  $s_c$  be the least index  $s$  such that  $a_s$  exceeds  $\sigma_c$  (assuming that the expansion of  $\Omega_f$  is not eventually constant and leaving the similar considerations for this case to the reader). Then the sum over all values  $2^{-f(n)}$  such that none of the increases  $d_0$  through  $d_s$

was due to  $n$  is at most  $2^{-|\sigma_c|}$ , hence all pairs of the form  $(f(n, s) - |\sigma_c| + 1, n)$  for such  $n$  and  $s$  where either  $s = 0$  or  $f(n, s)$  differs from  $f(n, s - 1)$  form a sequence of Kraft-Chaitin axioms, which is uniformly effective in  $c$  and  $\sigma_c$  relative to oracle  $A$ . Observe that by construction for each  $n$ , there is an axiom of the form  $(f(n) - |\sigma_c| + 1, n)$  and the sum of all terms  $2^{-k}$  over all axioms of the form  $(k, n)$  is less than  $2^{-f(n) - |\sigma_c|}$ .

Now consider a prefix-free Turing machine  $M$  with oracle  $A$  that given codes for  $c$  and  $\sigma_c$  and some other word  $p$  as input, first computes  $c$  and  $\sigma_c$ , then searches for  $s_c$ , and finally outputs the word that is coded by  $p$  according to the Kraft-Chaitin axioms for  $c$ , if such a word exists. If we let  $d$  be the coding constant for  $M$ , we have for all sufficiently large  $c$  and  $x$  that  $K^A(n) \leq 2 \log c + K^A(\sigma_c) + f(n) - |\sigma_c| + 1 + d \leq f(n) - c$ .  $\square$

As special cases of Theorem 2 we obtain the following results by Bienvenu and Downey [BD] and by Miller [M], where the former one is immediate and for the latter one it suffices to observe that the definition of the notion low for  $\Omega$  in terms of Chaitin's  $\Omega$  number

$$\Omega := \sum_{\{x: \mathbb{U}(x) \downarrow\}} 2^{-|x|}.$$

is equivalent to a definition in terms of  $\Omega_K$ .

**Corollary 1 (Bienvenu and Downey).** *A computable information content measure  $f$  is a Solovay function if and only if  $\Omega_f$  is Martin-Löf random.*

**Corollary 2 (Miller).** *A set  $A$  is weakly low if and only if  $A$  is low for  $\Omega$ .*

*Proof.* In order to see the latter result, it suffices to let  $f = K$  and to recall that for this choice of  $f$  the properties of  $A$  that occur in the two equivalent assertions in the conclusion of Theorem 2 coincide with the concepts weakly low and low for  $\Omega_K$ . But the latter property is equivalent to being low for  $\Omega$ , since for any set  $A$ , it is equivalent to require that some or that all left-computable Martin-Löf random set are Martin-Löf random relative to  $A$  [N, Proposition 8.8.1].  $\square$

By Corollary 1 and Theorem 1 it is immediate that the known Martin-Löf randomness of  $\Omega_K$  extends to the time-bounded case.

**Corollary 3.** *There is a constant  $c_0$  such that  $\Omega_{K^t} := \sum_{x \in \mathbb{N}} 2^{-K^t(x)}$  is Martin-Löf random for any computable function  $t$  where  $c_0 n \leq t(n)$  for almost all  $n$ .*

### 3 Solovay Functions and Jump-Traceability

In an attempt to define  $K$ -triviality without resorting to effective randomness or measure, Barmpalias, Downey and Greenberg [BDG] searched for characterizations of  $K$ -triviality via jump-traceability. They demonstrated that  $K$ -triviality is not implied by being  $h$ -jump-traceable for all computable functions  $h$  such that  $\sum_n 1/h(n)$  converges. Subsequently, the following question received some

attention: Can K-triviality be characterized by being  $g$ -jump traceable for all computable functions  $g$  such that  $\sum 2^{-g(n)}$  converges, that is, for all computable functions  $g$  that, up to an additive constant term, are upper bounds for  $K$ ?

We will now argue that Solovay functions can be used for a characterization of K-triviality in terms of jump traceability. However, we will not be able to completely avoid the notion of Kolmogorov complexity.

**Definition 4.** *A set  $A$  is K-trivial if  $K(A \upharpoonright n) \leq^+ K(n)$  for all  $n$ .*

**Definition 5.** *Let  $h : \mathbb{N} \rightarrow \mathbb{N}$  be a computable function. A set  $A$  is  $O(h(n))$ -jump-traceable if for every function  $\Phi$  partially computable in  $A$  there is a function  $h \in O(h(n))$  and a sequence  $(T_n)_{n \in \mathbb{N}}$  of uniformly c.e. finite sets, which is called a trace, such that for all  $n$*

$$|T_n| \leq h(n), \quad \text{and} \quad \Phi(n) \in T_n$$

for all  $n$  such that  $\Phi(n)$  is defined.

**Theorem 3.** *There is a constant  $c_0$  such that the following assertions are equivalent for any set  $A$ .*

- (i)  *$A$  is K-trivial.*
- (ii)  *$A$  is  $O(g(n) - K(n))$ -jump-traceable for every Solovay function  $g$ .*
- (iii)  *$A$  is  $O(K^t(n) - K(n))$ -jump-traceable for all computable functions  $t$  where  $c_0 n \leq t(n)$  for almost all  $n$ .*
- (iv)  *$A$  is  $O(K^t(n) - K(n))$ -jump-traceable for some computable function  $t$  where  $c_0 n \leq t(n)$  for almost all  $n$ .*

*Proof.* The implication (ii) $\Rightarrow$ (iii) is immediate by Theorem 1, and the implication (iii) $\Rightarrow$ (iv) is trivially true. So it suffices to show the implications (i) $\Rightarrow$ (ii) and (iv) $\Rightarrow$ (i), where due to space considerations we only sketch the corresponding proofs.

First, let  $A$  be K-trivial and let  $\Phi^A$  be any partially  $A$ -computable function. Let  $\langle \cdot, \cdot \rangle$  be some standard effective pairing function. Since  $A$  is K-trivial and hence low for  $K$ , we have

$$K(\langle n, \Phi^A(n) \rangle) =^+ K^A(\langle n, \Phi^A(n) \rangle) =^+ K^A(n) =^+ K(n) ,$$

whenever  $\Phi^A(n)$  is defined. Observe that the constant that is implicit in the relation  $=^+$  depends only on  $A$  in the case of the first and last relation symbol, but depends also on  $\Phi$  in case of the middle one.

By the coding theorem there can be at most constantly many pairs of the form  $(n, y)$  such that  $K(n, y)$  and  $K(n)$  differ at most by a constant, and given  $n$ ,  $K(n)$  and the constant, we can enumerate all such pairs. But then given a Solovay function  $g$ , for given  $n$  we can enumerate at most  $g(n) - K(n) + 1$  possible values for  $K(n)$  and for each such value at most constantly many pairs  $(n, y)$  such that some  $y$  is equal to  $\Phi^A(n)$ , hence  $A$  is  $O(g(n) - K(n))$ -jump-traceable.

Next, let  $c_0$  be the constant from Theorem 1 and let  $t$  be a computable time bound such that (iv) is true for this value of  $c_0$ . Then  $K^t$  is a Solovay function by choice of  $c_0$ .

Recall the tripling function  $[\cdot, \cdot, \cdot]$  and the concept of a Solovay triple  $[s, \sigma, n]$  from the proof of Theorem 1, and define a partial  $A$ -computable function  $\Phi$  that maps any Solovay triple  $[s, \sigma, n]$  to  $A \upharpoonright n$ . Then given an optimal code  $\sigma$  for  $n$ , one can compute the corresponding Solovay triple  $z = [s, \sigma, n]$ , where then  $K^t(z)$  and  $K(z)$  differ only by a constant, hence the trace of  $\Phi^A$  at  $z$  has constant size and contains the value  $A \upharpoonright n$ , i.e., we have  $K(A \upharpoonright n) \leq^+ |\sigma| = K(n)$ , hence  $A$  is  $K$ -trivial.  $\square$

### 4 Time Bounded Kolmogorov Complexity and Strong Depth

The initial segments of a c.e. set  $A$  have small Kolmogorov complexity, by Barzdins' lemma [DH] it holds for all  $m$  that

$$C(A \upharpoonright m \mid m) \leq^+ \log m \quad \text{and} \quad C(A \upharpoonright m) \leq^+ 2 \log m.$$

Furthermore, there are infinitely many initial segments that have considerably smaller complexity. The corresponding observation in the following remark is extremely easy, but apparently went unnoticed so far and, in particular, improves on corresponding statements in the literature [DH, Lemma attributed to Solovay in Chapter 14].

*Remark 2.* Let  $A$  be a c.e. set. Then there is a constant  $c$  such that for infinitely many  $m$  it holds that

$$C(A \upharpoonright m \mid m) \leq c, \quad C(A \upharpoonright m) \leq^+ C(m) + c, \quad \text{and} \quad C(A \upharpoonright m) \leq \log m + c.$$

For a proof, it suffices to fix an effective enumeration of  $A$  and to observe that there are infinitely many  $m \in A$  such that  $m$  is enumerated after all numbers  $n \leq m$  that are in  $A$ , i.e., when knowing  $m$  one can simulate the enumeration until  $m$  appears, at which point one then knows  $A \upharpoonright m$ .

Barzdins [Ba] states that there are c.e. sets with high time-bounded Kolmogorov complexity, and the following lemma generalizes this in so far as such sets can be found in every high Turing degree.

**Lemma 1.** *For any high set  $A$  there is a set  $B$  where  $A =_T B$  such that for every computable time bound  $t$  there is a constant  $c_t > 0$  where*

$$C^t(B \upharpoonright m) \geq^+ c_t \cdot m \quad \text{and} \quad C(B \upharpoonright m) \leq^+ 2 \log m.$$

*Moreover, if  $A$  is c.e.,  $B$  can be chosen to be c.e. as well.*

*Proof.* Let  $A$  be any high set. We will construct a Turing-equivalent set  $B$  as required. Recall the following equivalent characterization of a set  $A$  being high:

there is a function  $g$  computable in  $A$  that majorizes any computable function  $f$ , i.e.,  $f(n) \leq g(n)$  for almost all  $n$ . Fix such a function  $g$ , and observe that in case  $A$  is c.e., we can assume that  $g$  can be effectively approximated from below. Otherwise we may replace  $g$  with the function  $g'$  defined as follows. Let  $M_g$  be an oracle Turing machine that computes  $g$  if supplied with oracle  $A$ . For all  $n$ , let

$$\tilde{g}(n, s) := \max\{M_g^{A_i}(n) \mid i \leq s\},$$

where  $A_i$  is the approximation to  $A$  after  $i$  steps of enumeration, and let  $g'(n) := \lim_{s \rightarrow \infty} \tilde{g}(n, s)$ . We have  $g(n) \leq g'(n)$  for all  $n$  and by construction,  $g'$  can be effectively approximated from below.

Partition  $\mathbb{N}$  into consecutive intervals  $I_0, I_1, \dots$  where interval  $I_j$  has length  $2^j$  and let  $m_j = \max I_j$ . By abuse of notation, let  $t_0, t_1, \dots$  be an effective enumeration of all partial computable functions. Observe that it is sufficient to ensure that the assertion in the theorem is true for all  $t = t_i$  such that  $t_i$  is computable, nondecreasing and unbounded. Assign the time bounds to the intervals  $I_0, I_1, \dots$  such that  $t_0$  will be assigned to every second interval including the first one,  $t_1$  to every second interval including the first one of the *remaining* intervals, and so on for  $t_2, t_3, \dots$ , and note that this way  $t_i$  will be assigned to every  $2^{i+1}$ -th interval.

We construct a set  $B$  as required. In order to code  $A$  into  $B$ , for all  $j$  let  $B(m_j) = A(j)$ , while the remaining bits of  $B$  are specified as follows. Fix any interval  $I_j$  and assume that this interval is assigned to  $t = t_i$ . Let  $B$  have empty intersection with  $I_j \setminus \{m_j\}$  in case the computation of  $t(m_j)$  requires more than  $g(j)$  steps. Otherwise, run all codes of length at most  $|I_j| - 2$  on the universal machine  $\mathbb{V}$  for  $2t(m_j)$  steps each, and let  $w_j$  be the least word of length  $|I_j| - 1$  that is not output by any of these computations, hence  $C^{2t}(w_j) \geq |w_j| - 1$ . Let the restriction of  $B$  to the first  $|w_j|$  places in  $I_j$  be equal to  $w_j$ .

Now let  $v_j$  be the initial segment of  $B$  of length  $m_j + 1$ , i.e., up to and including  $I_j$ . In case  $t = t_i$  is computable, nondecreasing and unbounded, for almost all intervals  $I_j$  assigned to  $t$ , we have  $C^t(v_j) > |v_j|/3$ , because otherwise, for some appropriate constant  $c$ , the corresponding codes would yield for almost all  $j$  that  $C^{2t}(w_j) \leq |v_j|/3 + c \leq |I_j| - 2$ . Furthermore, by construction for every such  $t$  there is a constant  $c_t > 0$  such that for almost all  $m$ , there is some interval  $I_j$  assigned to  $t$  such that  $m_j \leq m$  and  $c_t m \leq m_j/4$ , hence for almost all  $m$  the initial segment of  $B$  up to  $m$  cannot have Kolmogorov complexity of less than  $c_t m$ .

We omit the routine proof that  $A$  and  $B$  are Turing-equivalent and that if  $A$  was c.e., then  $B$  is c.e. as well, where for the latter fact we need the assumption that  $g$  can be effectively approximated from below.

Finally, to see that  $C(B \upharpoonright m) \leq^+ 2 \log m$ , notice that in order to determine  $B \upharpoonright m$  without time bounds it is enough to know on which of the intervals  $I_j$  the assigned time bounds  $t_i$  terminate before their computation is canceled by  $g$ , which requires one bit per interval, plus another one describing the bit of  $A$  coded into  $B$  at the end of each interval.  $\square$

**Lemma 2.** *Every high degree contains for every computable, nondecreasing and unbounded function  $h$  a set  $B$  such that for every computable time bound  $t$  and almost all  $m$ ,*

$$C^t(B \upharpoonright m) \geq^+ \frac{1}{4}m \quad \text{and} \quad C(B \upharpoonright m) \leq h(m) \cdot \log m.$$

*Proof.* The argument is similar to the proof of Lemma 1, but now, when considering interval  $I_j$ , we diagonalize against the largest running time among  $t_0(m_j), \dots, t_{h(j)-2}(m_j)$  such that the computation of this value requires not more than  $g(j)$  steps. This way we ensure – for any computable time bound  $t$  – that at the end of almost all intervals  $I_j$  compression by a factor of at most  $1/2$  is possible, and that within interval  $I_j$ , we have compressibility by a factor of at most  $1/4$ , up to a constant additive term, because interval  $I_{j-1}$  was compressible by a factor of at most  $1/2$ .  $\square$

Kummer's gap theorem asserts that any array noncomputable c.e. Turing degree contains a c.e. set  $A$  such that there are infinitely many  $m$  such that  $C(A \upharpoonright m) \geq 2 \log m$ , whereas all c.e. sets in an array computable Turing degree satisfy  $C(A \upharpoonright m) \leq (1 + \varepsilon) \log m$  for all  $\varepsilon > 0$  and almost all  $m$ . Similarly, Theorem 4, the main result of this section, asserts a dichotomy for the time-bounded complexity of initial segments between high and nonhigh sets.

**Theorem 4.** *Let  $A$  be any c.e. set.*

- (i) *If  $A$  is high, then there is a c.e. set  $B$  with  $B =_T A$  such that for every computable time bound  $t$  there is a constant  $c_t > 0$  such that for all  $m$ , it holds that  $C^t(B \upharpoonright m) \geq c_t \cdot m$ .*
- (ii) *If  $A$  is not high, then there is a computable time bound  $t$  such that  $C^t(A \upharpoonright m) \leq^+ \log m$ .*

*Proof.* The first assertion is immediate from Lemma 1. In order to demonstrate the second assertion, it suffices to observe that for the modulus of convergence  $s$  of the c.e. set  $A$  there is a computable function  $f$  such that  $s(m) \leq f(m)$  for infinitely many  $m$ .  $\square$

As another easy consequence of Lemma 1, we get an alternative proof of the result due to Juedes, Lathrop and Lutz [JLL] that every high degree contains a strongly deep set. We omit details due to lack of space.

## References

- [BDG] Barmpalias, G., Downey, R., Greenberg, N.: K-trivial degrees and the jump-traceability hierarchy. Proc. Amer. Math. Soc.; posted on January 22, 2009, PII S 0002-9939(09)09761-5 (to appear in print)
- [Ba] Barzdins, J.M.: Complexity of programs to determine whether natural numbers not greater than  $n$  belong to a computably enumerable set. Soviet Math. Dokl. 9, 1251–1254 (1968)

- [BD] Bienvenu, L., Downey, R.: Kolmogorov Complexity and Solovay Functions. In: Symposium on Theoretical Aspects of Computer Science , STACS 2009, pp. 147–158 (2009), <http://drops.dagstuhl.de/opus/volltexte/2009/1810>
- [DH] Downey, R., Hirschfeldt, D.: Algorithmic Randomness (manuscript, 2009)
- [JLL] Juedes, D., Lathrop, J., Lutz, J.: Computational depth and reducibility. *Theor. Comput. Sci.* 132(2), 37–70 (1994)
- [K] Kummer, M.: Kolmogorov complexity and instance complexity of recursively enumerable sets. *SIAM Journal on Computing* 25, 1123–1143 (1996)
- [KS] Kučera, A., Slaman, T.: Randomness and recursive enumerability. *SIAM Journal on Computing* 31, 199–211 (2001)
- [LV] Li, M., Vitányi, P.: *An Introduction to Kolmogorov Complexity and Its Applications*, 3rd edn. Springer, Heidelberg (2008)
- [M] Miller, J.S.: *The K-degrees, low for K-degrees and weakly low for K-degrees*. *Notre Dame Journal of Formal Logic* (to appear)
- [N] Nies, A.: *Computability and Randomness*. Oxford University Press, Oxford (2008)
- [S] Solovay, R.: Draft of a paper (or series of papers) on Chaitin’s work. Unpublished notes, 215 pages (1975)