# ANALOGUES OF CHAITIN'S OMEGA IN THE COMPUTABLY ENUMERABLE SETS

G. BARMPALIAS, R. HÖLZL, A.E.M. LEWIS, AND W. MERKLE

Abstract. We show that there are computably enumerable (c.e.) sets with maximum initial segment Kolmogorov complexity amongst all c.e. sets (with respect to both the plain and the prefix-free version of Kolmogorov complexity). These c.e. sets belong to the weak truth table degree of the halting problem, but not every weak truth table complete c.e. set has maximum initial segment Kolmogorov complexity. Moreover, every c.e. set with maximum initial segment prefix-free complexity is the disjoint union of two c.e. sets with the same property; and is also the disjoint union of two c.e. sets of lesser initial segment complexity.

## 1. Introduction

Kolmogorov complexity measures the complexity of a finite sequence in terms of the shortest program that can generate it. It may also be used in order to study the initial segment complexity of infinite sequences, and it is this approach that led to the definition of random sequences in [Lev73, Cha75]. Measures of relative initial segment complexity were initially introduced for the class of computably enumerable (c.e.) reals (i.e. reals that are the limits of increasing computable sequences of rationals) and were used in order to characterize Chaitin's $\Omega$ numbers as the c.e. reals with maximum initial segment complexity. In this note we are concerned with the initial segment complexity of c.e. sets. We discover a class of c.e. sets of maximum initial segment complexity and study some of its properties. These c.e. sets may be seen as analogues of Chaitin's $\Omega$ numbers in the class of c.e. sets.

In Section 1.1 we review the measures that have been used in the literature in order to classify classes of reals according to their initial segment complexity. In Section 1.2 we give an account of the known properties concerning the initial

---

*Last revision:* January 5, 2013.

*Key words and phrases.* Kolmogorov complexity, computably enumerable sets, completeness.

This research was partially done during the Dagstuhl Seminar on Computability, Complexity and Randomness (12021) and the the programme 'Semantics and Syntax' in the Isaac Newton Institute for the Mathematical Sciences, Cambridge U.K. Barmpalias was supported by the *Research fund for international young scientists* number 611501-10168 from the National Natural Science Foundation of China, and an *International Young Scientist Fellowship* number 2010-Y2GB03 from the Chinese Academy of Sciences; partial support was also received from the project *Network Algorithms and Digital Information* number ISCAS2010-01 from the Institute of Software, Chinese Academy of Sciences. Hölzl was supported by a Feodor Lynen postdoctoral research fellowship by the Alexander von Humboldt Foundation. Lewis was supported by a Royal Society University Research Fellowship. This work was partially supported by Bulgarian National Science Fund under contract D002-258/18.12.08. We are indebted to the referees, and in particular one of the referees who suggested Theorem 2.7 and simplified the proof of Theorem 2.4.

segment complexity of c.e. sets. In Section 1.3 we give an outline of our results, which are presented in detail in the main part of this note.

1.1. **Measures of relative initial segment complexity.** One of the earliest measures for comparing the initial segment complexity of *reals* (which we identify with their binary expansion) was introduced and studied in [Sol75]. It is known as the 'Solovay reducibility', often denoted by $\leq_S$, and for c.e. reals it essentially measures the hardness of approximation 'from below'. It is a preorder and it induces a partially ordered degree structure that is known as the Solovay degrees. In a series of papers [Sol75, CHKW01, KS01] it was shown that the random c.e. reals are exactly the reals in the greatest Solovay degree, and they coincide with the halting probabilities of universal prefix-free machines. This structure was studied further (see [DH10, Section 9.5] for an overview) and was generally accepted as an adequate measure for classifying initial segment complexities for the class of c.e. reals. A number of related measures were introduced in [DHL04] with the hope of providing measures of relative complexity for different classes of reals. Let $K_M$ denote the Kolmogorov complexity function with respect to the Turing machine $M$ (i.e. $K_M(\sigma)$ is the length of the shortest string $\tau$ such that $M(\tau) = \sigma$, and $\infty$ if this does not exist). Let $K = K_U$ where $U$ is a fixed optimal universal prefix-free machine and let $C = K_V$ where $V$ is a fixed optimal universal (plain) Turing machine. Also, let $C(\sigma|\tau)$ denote the Kolmogorov complexity of $\sigma$ relative to $\tau$ (i.e. when $\tau$ is given as an oracle in the underlying machine that describes $\sigma$). A real $X$ is called random if $\exists c \forall n,\ K(X\restriction_n) \geq n - c$. Perhaps the most straightforward measure of relative initial segment complexity is $\leq_K$ (already implicit in [Sol75]).

$$X \leq_K Y \overset{\text{def}}{\iff} \exists c \forall n\ (K(X\restriction_n) \leq K(Y\restriction_n) + c). \tag{1.1}$$

We may express the fact that $X \leq_K Y$ simply by saying that the prefix-free initial segment complexity of $X$ is less than (or equal to) the prefix-free initial segment complexity of $Y$. The plain complexity version $\leq_C$ of the above relation is defined analogously. These preorders induce the $K$-degrees and the $C$-degrees respectively, which have received a certain amount of attention (see [DH10, Section 9.7]). We note that $\leq_S$ is contained in $\leq_K$ and so the $K$-degrees of c.e. reals have a largest element that contains the random c.e. reals. The main proposal for an alternative to Solovay reducibility that applies to more general classes of reals was the *relative $K$-reducibility* (in symbols, $rK$), which is defined by

$$X \leq_{rK} Y \overset{\text{def}}{\iff} \exists c \forall n\ (K(X\restriction_n \mid Y\restriction_n) \leq c). \tag{1.2}$$

Note that $X \leq_{rK} Y$ can be defined equivalently using plain complexity, by the relation $\exists c \forall n (C(X\restriction_n \mid Y\restriction_n) \leq c)$. This follows from the basic relations between plain and prefix-free complexity, namely the fact that there exists a constant $d$ such that $C(\sigma|\tau) \leq K(\sigma|\tau) + d$ and $K(\sigma|\tau) \leq 2C(\sigma|\tau) + d$ for all strings $\sigma, \tau$. It is not hard to see that the relation $X \leq_{rK} Y$ is equivalent to (1.3).

(1.3) There exists a partial computable function $f : 2^{<\omega} \times \omega \to 2^{<\omega}$ and a constant $c$ such that $\forall n \exists j < c\ (f(Y\restriction_n, j)\downarrow = X\restriction_n)$.

Here $\omega$ denotes the set of natural numbers and $2^{<\omega}$ denotes the set of finite binary strings. This shows that $X \leq_{rK} Y$ implies $X \leq_K Y$ and $X \leq_C Y$. Moreover (as observed in [DHL04]) $X \leq_{rK} Y$ implies $X \leq_T Y$ (where $\leq_T$ is the Turing reducibility). In [MS07] it was observed that $X \leq_C Y$ implies $X \leq_T Y$, provided that $Y$ is a subset of $\{2^{2^n} \mid n \in \omega\}$.

1.2. **The initial segment complexity of c.e. sets.** By [Bar68], if $A$ is a c.e. set then $\exists c \forall n,\ C(A \restriction_n) \leq 2\log n + c$ (here and throughout this paper, $\log n$ denotes the largest integer which is less than or equal to the logarithm of $n$ on base 2); on the other hand, there are c.e. sets $B$ such that $\forall n\ C\ (B \restriction_n) \geq \log n - b$ for some constant $b$. Each of these observations lead to a more informative view about c.e. sets with complicated initial segments. The first is from [Kum96] and is known as the *Kummer dichotomy*. It says that every member in a certain class of c.e. Turing degrees that is known as the *array non-computable degrees* contains a c.e. set $A$ such that $C(A \restriction_n) \geq 2\log n - c$ for infinitely many $n$ and some constant $c$; on the other hand if the degree of a c.e. set $B$ is not in that class and $f$ is any computable order (i.e. nondecreasing unbounded function) then $\exists b \forall n,\ C(B \restriction_n) \leq \log n + f(n) + b$. The second is from [Kan69] and [KHMS11] and characterizes the c.e. sets $A$ such that $\forall n,\ C(A \restriction_n) \geq f(n)$ for a computable order $f$, as the weak truth table complete c.e. sets (i.e. the sets that compute the halting problem with a computable bound on their use in the computation). These are also called *complex sets*.

Further research on this topic concerns the behavior of the measures of complexity that we discussed in Section 1.1 on the class of c.e. sets. In [Bar05] it was shown that in the Solovay degrees of c.e. sets there are pairs with no upper bound; in particular, there is no maximum degree. In [Bar11b] it was shown that there are no minimal pairs in the structure of the $K$-degrees of c.e. sets. This gave an elementary property that distinguishes this structure from the $C$-degrees, the $rK$-degrees and the Solovay degrees of c.e. sets. A number of other features in the $K$-degrees and the $C$-degrees degrees of c.e. sets (including splitting theorems and cone avoidance arguments) were shown in [Ste11, Chapter 2] (see [Bar11a, Section 5]) and [BL11, Section 6] by emulating the corresponding arguments in the c.e. Turing degrees.

1.3. **Our results.** Perhaps the most basic question concerning the relative initial segment complexity of c.e. sets is whether there exist c.e. sets that are more complex (modulo a constant) than any other c.e. set. In view of the Kummer dichotomy and the behavior of the c.e. sets in the Solovay degrees, one would guess a negative answer for any of the measures of relative complexity of Section 1.1. In Section 2 we show that, surprisingly, there are complete (i.e. maximum) elements in the partial orders of the $rK$-degrees, the $C$-degrees and the $K$-degrees of c.e. sets.

> There are c.e. sets of maximum initial segment complexity amongst all c.e. sets.

We also show that they form a proper subclass of the complex sets of [Kan69], [KHMS11] and provide a characterisation of this class for the case of the $rK$-degrees. In Section 3 we show some splitting properties of the c.e. sets in the $rK$-degrees, the $C$-degrees and the $K$-degrees. In particular, every c.e. set can be split into two disjoint c.e. sets of the same $K$-degree, while this fails in the cases of the $rK$-degrees and the $C$-degrees. The former result is rather surprising and can be seen as a formal expression of the following statement.

> Every c.e. set is the disjoint union of two c.e. sets of the same
> initial segment prefix-free complexity.

As a consequence, every c.e. complex set is the disjoint union of two complex c.e. sets and an analogous result holds for the $K$-complete c.e. sets.

It remains to be seen if the $K$-degrees of the c.e. sets are dense; by the splitting theorems of [Ste11, Chapter 2], [Bar11a, Section 5] they are downward dense. A relevant open question is whether every pair of c.e. sets has a least upper bound in the c.e. $K$-degrees.

## 2. Maximum initial segment complexity in the c.e. sets

We show that there exists a c.e. set $B$ such that $W \leq_{rK} B$ for all c.e. sets $W$. According to the discussion of Section 1.1 this set will also satisfy $W \leq_C B$ and $W \leq_K B$ for all c.e. sets $W$. We will use a way of ensuring one set is $rK$-reducible to another, which is particular to the class of c.e. sets. Condition (2.1) of Lemma 2.1 says that for every $\ell$ and for every $c$ enumerations into $A\restriction_\ell$, at least one enumeration into $B\restriction_\ell$ occurs.

**Lemma 2.1** ($rK$-reductions for c.e. sets)**.** *Let $A, B$ be c.e. sets with enumerations satisfying the following property:*

$$\text{(2.1)}\qquad \textit{There exists } c \textit{ such that for every } \ell \textit{ and } s<t, \textit{ if } |A[t]\restriction_\ell - A[s]\restriction_\ell| \geq c \textit{ then } |B[t]\restriction_\ell| > |B[s]\restriction_\ell|.$$

*Then $A \leq_{rK} B$.*

**Proof.** It suffices to define a partial computable function $f$ that meets (1.3). The definition of $f$ is as follows. For each $\sigma$ and $j$, wait for a stage $s$ such that $B[s]\restriction_{|\sigma|}= \sigma$ and the remainder of $|A[s]\restriction_{|\sigma|}|$ divided by $c$ is $j$. Then let $f(\sigma, j) = A[s]\restriction_{|\sigma|}$.

For the verification, first we show that for each $\ell$, if $f(B\restriction_\ell, j)$ is defined and $j$ is the remainder of $|A\restriction_\ell|$ divided by $c$ then it equals $A\restriction_\ell$. Indeed, suppose that the definition occurred at stage $s$. If $A[s]\restriction_\ell$ is not correct, there must occur at least $c$ enumerations into $A\restriction_\ell$ after stage $s$. But according to condition (2.1) this means that $B[s]\restriction_\ell$ is not a prefix of $B$, a contradiction. Finally for each $\ell$ it is clear that $f(B\restriction_\ell, j)$ will be defined if $j$ is the remainder of $|A\restriction_\ell|$ divided by $c$. ☐

We may now use Lemma 2.1 in order to prove that there are $rK$-complete c.e. sets.

**Theorem 2.2** ($rK$-completeness)**.** *There exists an $rK$-complete c.e. set $B$. In particular, $W \leq_{rK} B$ for all c.e. sets $W$; moreover the reductions are uniform in the indices of the c.e. sets.*

**Proof.** We make use of Lemma 2.1 in order to construct the required reductions. Without loss of generality, we only deal with c.e. sets $W$ such that $W(0) = W(1) = 0$. For the duration of this proof we let $(W_e)$ be an effective enumeration of all c.e. sets that satisfy this condition, such that at any stage $s$ at most one number is enumerated in $\cup_{e<s} W_e$. As usual (by a standard convention), $n \in W_e[s]$ implies $n < s$ and $e < s$. Note that, given the nature of the $rK$-reducibility, a uniform sequence of reductions for this restricted class of sets is sufficient for the proof of the theorem. Let $I_e(i) = [2^i, 2^{i+1}) \cap W_e$. Clearly, this family of sets is uniformly c.e. and $|I_e(i)| \leq 2^i$ for all $e, i$. We define $B$ by enumerating it during the stages of the universal enumeration of $(W_e)$ as follows. For each $i > 0$ and each $2^{e+3}$ times that a number is enumerated into $\cup_{t\leq i} I_e(t)$, we enumerate into $B$ the least number in $[2^{i-1}, 2^i)$ which is currently out of $B$. Formally, at stage $s+1$ for each $e$ and $i > 0$ such that

$$2^{e+3} \mid \Big|\bigcup_{t\leq i} I_e(t)[s+1]\Big| \quad \text{and} \quad \Big|\bigcup_{t\leq i} I_e(t)[s+1]\Big| > \Big|\bigcup_{t\leq i} I_e(t)[s]\Big|$$

we enumerate the least element of $[2^{i-1}, 2^i) - B[s]$ into $B[s+1]$.

First, note that the enumeration of $B$ is computable, although there is an unbounded quantifier on $e, i$ in each step. This is because $I_e(i)[s] = \emptyset$ whenever $e \geq s$ or $i \geq s$. Second, fix $i > 0$ and note that at each stage at most one number from $[2^{i-1}, 2^i)$ is requested by the construction to be enumerated in $B$. Moreover every such request corresponds to some (possibly many) $e$. We claim that for each $e$, there can be at most $2^{i+1}/2^{e+3} = 2^{i-e-2}$ corresponding requests for enumeration into $B \cap [2^{i-1}, 2^i)$. This follows from the construction and the fact that $|\cup_{t\leq i} I_e(t)| \leq 2^{i+1}$. Hence, overall there can be at most $\sum_e 2^{i-e-2} = 2^{i-1}$ requests for enumeration into $B \cap [2^{i-1}, 2^i)$. This means that if at some stage $s+1$ the construction requests an enumeration of a number in $[2^{i-1}, 2^i) - B[s]$ into $B$, we have $[2^{i-1}, 2^i) - B[s] \neq \emptyset$ (i.e. such an enumeration will occur).

Finally, fix $e$ and observe that condition (2.1) of Lemma 2.1 holds for $c = 2^{e+3}$ and $A = W_e$. Indeed, pick *any* $\ell$ and let $i$ be the least number such that $\ell \leq 2^{i+1}$. Then enumerations in $W_e \upharpoonright_\ell$ are also enumerations in $\cup_{t\leq i} I_e(t)$. Hence during every $2^{e+3}$ enumerations in $W_e \upharpoonright_\ell$, an enumeration will occur in $B \cap [2^{i-1}, 2^i)$, which is also an enumeration in $B \upharpoonright_\ell$. □

The following consequence is immediate, in view of the discussion of Section 1.1.

**Corollary 2.3.** *There exists a $\leq_C$-complete and $\leq_K$-complete c.e. set.*

Recall that an order is a nondecreasing unbounded function, and that a set $X$ is called *complex* if there is a computable order $f$ such that $\forall n,\ C(A\upharpoonright_n) \geq f(n)$. This is equivalent to the condition that $\forall n,\ K(A\upharpoonright_n) \geq g(n)$ for some computable order $g$. In [Kan69] and [KHMS11] it was shown that a c.e. set is complex if and only if it is in the same weak truth table degree as the halting problem. We clarify the relation between complex sets and $K$-complete (or $C$-complete) sets. Moreover, we observe that there are many-one complete c.e. sets which are not $\leq_K$-complete or $\leq_C$-complete. The proof of the following theorem was simplified by one of the referees.

**Theorem 2.4.** *Every $\leq_C$-complete and every $\leq_K$-complete c.e. set is in the weak truth table degree of the halting problem. The converse does not hold.*

**Proof.** Clearly, every $\leq_K$-complete and every $\leq_C$-complete c.e. set is complex. Hence they are in the weak truth table degree of the halting problem.

In order to show that the converse does not hold, it suffices to consider the set $B = \{2^{2^x} \mid x \in \emptyset'\}$. This set is clearly in the weak truth table degree of the halting problem but below each number $n$ it has at most $\log\log n$ elements. Hence its plain and prefix-free initial segment complexities are bounded above by $\log n + 2\log\log(n)$ and $\log n + 4\log\log(n)$ (modulo a constant) respectively. By the discussion at the beginning of Section 1.2 such a set is neither $C$-complete nor $K$-complete. □

The above classes of c.e. sets with maximum initial segment complexity appear to be new. Hence any characterisation of them in terms of notions from classical computability theory is desirable. We take a step in this direction by providing such a characterisation for the case of $rK$-complete c.e. sets. This investigation will also produce natural examples of sets in this class. The following observation says that c.e. sets can be compressed by any multiplicative factor and is crucial to our argument.

**Lemma 2.5.** *Let $A$ be a c.e. set and $a, b \in \omega$. Then there exists a c.e. set $D$ and a constant $c$ such that $C(A\upharpoonright_{(an+b)} \mid D\upharpoonright_n) \le c$ for all $n$.*

**Proof.** For every set $A$ it is clear that $C(A\upharpoonright_{(n+b)} \mid A\upharpoonright_n)$ is bounded by a constant. Hence (by the transitivity of the $rK$ reductions) we may assume that $b = 0$ in the statement of the lemma, without loss of generality. This argument is very similar to the construction in the proof of Theorem 2.2, so we write it a bit more concisely. Let $I_j = [2^j, 2^{j+1})$, fix a computable enumeration $(A[s])$ of $A$ and define the c.e. set $D$ dynamically with respect to the stages $s$ as follows. For every $j > 0$, each time we observe $8 \cdot a$ new enumerations of numbers $< a \cdot 2^{j+1}$ into $A$, we enumerate into $D$ the largest number of $[2^{j-1}, 2^j)$ which is not yet in $D$.

It is clear that there will always be space in $D$ for such enumerations, so the construction of $D$ is well defined. Moreover, for each $n$ we may describe $A\upharpoonright_{an}$ using $D\upharpoonright_n$ as follows. Find a stage $s$ such that $D\upharpoonright_n = D[s]\upharpoonright_n$. According to the construction, it is not possible that $8 \cdot a$ numbers will be enumerated into $A\upharpoonright_{an}$ at stages after stage $s$. Indeed, this would produce an enumeration of a number $\le n$ into $D$ (which contradicts the choice of $s$). Therefore we may describe $A\upharpoonright_{an}$ by a string that indicates the number of the subsequent enumerations into $A\upharpoonright_{an}$. Since this number is less than $8 \cdot a$, there is a fixed upper bound on the length of such a string, which is independent of $n$. This shows that there is a constant $c$ such that $C(A\upharpoonright_{an} \mid D\upharpoonright_n) \le c$ for all $n$. □

An interesting consequence of this observation is that $rK$-complete sets have the following self-compression feature. Given a rational number $q$, let $\lfloor q \rfloor$ denote the largest integer $\le q$.

**Corollary 2.6.** *If $a, b \in \omega$ and $B$ is $rK$-complete then both $C(B\upharpoonright_{(an+b)} \mid B\upharpoonright_n)$ and $C(B\upharpoonright_n \mid B\upharpoonright_{\lfloor n/a \rfloor})$ are bounded above by a constant.*

**Proof.** The first clause is a direct consequence of Lemma 2.5. The second clause follows from the first one, given that two successive multiples of $a$ can differ by at most a constant (namely, $a$). □

We may now state and prove the characterisation of the $rK$-complete c.e. sets in terms of Turing completeness with linear use of the oracle. The following result also shows that the $rK$-complete c.e. sets coincide with the c.e. sets whose plain initial segment complexity is bounded below by $\log n$ (these were discussed in the beginning of Section 1.2).

Let $D \le_{\text{lin}} G$ denote the fact that set $D$ is computable from set $G$ with a Turing reduction that has a linear use function. It is not hard to see that there are complete c.e. sets with respect to $\le_{\text{lin}}$ i.e. c.e. sets $A$ such that $W \le_{\text{lin}} A$ for all c.e. sets $W$. This can be seen if we choose a pairing function which is linear in the second argument, for example (see [Odi89, Section I.1] for a discussion of various pairing functions)

$$\langle n, m \rangle = 2^n(2m+1) - 1.$$

In this case the set $\{\langle e, m \rangle \mid m \in W_e\}$ is $\le_{\text{lin}}$-complete (i.e. complete with respect to $\le_{\text{lin}}$). In particular, certain choices of arithmetization guarantee that the standard form of the halting problem is $\le_{\text{lin}}$-complete. For the following proof recall that if $(M_e)$ is an effective sequence of all Turing machines then the standard universal machine U (for the purposes of plain complexity) is given by $U(0^e 1\sigma) = M_e(\sigma)$ (we call this the 'standard encoding of the universal machine'). Moreover $C_M$ denotes the plain complexity relative to machine $M$. The following theorem was suggested by one of the referees.

**Theorem 2.7.** *Given any c.e. set $A$, the following clauses are equivalent:*

(a) *$A$ is $rK$-complete;*
(b) *$A$ satisfies $C(A\upharpoonright_n) \geq \log n - c$ for some constant $c$ and all $n$;*
(c) *$A$ is linear-complete.*

*Here 'linear-complete' means '$\leq_{lin}$-complete'.*

**Proof.** Let us fix a c.e. set $A$. Since there are sets $D$ such that $C(D\upharpoonright_n) \geq \log n - c$ for some constant $c$ and all $n$, clause (a) implies clause (b). For (b)⇒(c) assume that $A$ meets clause (b) and let $W$ be a c.e. set. It suffices to show that $W \leq_{\text{lin}} A$. In order to construct this reduction we define a Turing machine $M$ (which we think of as a function $2^{<\omega} \to 2^{<\omega}$), and by the recursion theorem we may use its index $e$ in its own definition. We may build $M$ in terms of a c.e. set of requests $(\sigma, n)$ (requesting the description of string $\sigma$ with a string of length $n$) as long as we ensure that for each $n$ there are at most $2^n$ requests with the second component equal to $n$ (e.g. see [Nie09, Section 2.1]). This request set $S$ is defined as follows: at stage $s+1$, if $n \in W[s+1]-W[s]$ then enumerate $(A[s+1]\upharpoonright_{n2^{c+e+2}}, \log n)$ into $S$. Clearly for each $m$ we enumerate at most $2^m$ requests with the second component equal to $m$. Hence if $n \in W[s+1] - W[s]$ then $C_M(A[s+1]\upharpoonright_{n2^{c+e+2}}) = \log n$. We may now demonstrate how to compute $W(n)$ via queries to $A\upharpoonright_{2^{c+e+3}}$. It suffices to show that $n$ may only be enumerated into $W$ in the stages up to when the approximation to $A\upharpoonright_{2^{c+e+3}}$ reaches a limit. Indeed, otherwise the definition of $M$ would imply $C_M(A\upharpoonright_{n2^{c+e+2}}) = \log n$ which means (by the standard encoding of the universal machine) that

$$C(A\upharpoonright_{n2^{c+e+2}}) \leq \log n + e + 1.$$

Since $\log n + e + 1 = \log(n2^{c+e+2}) - c - 1$ this contradicts clause (b). Hence this reduction witnesses the fact that $W \leq_{\text{lin}} A$.

It remains to show that (c) implies (a), so assume that $A$ meets clause (c). Let $B$ be an $rK$-complete c.e. set. Then there exists a constant $a$ such that $A$ computes $B$ via a Turing reduction with use function $n \mapsto an$. Hence $C(B\upharpoonright_n \mid A\upharpoonright_{an})$ is bounded above by a constant. This means that $C(B\upharpoonright_{\lfloor n/a \rfloor} \mid A\upharpoonright_n)$ is bounded above by a constant. By Corollary 2.6 it follows that $C(B\upharpoonright_n \mid A\upharpoonright_n)$ is bounded above by a constant, so that $B \leq_{rK} A$ and $A$ is $rK$-complete. □

We conclude this section with a couple of remarks. Theorem 2.7 provides a number of natural examples of $rK$-complete sets, including many versions of the halting problem. Moreover Kolmogorov's set of non-random strings (i.e. strings $\sigma$ such that $C(\sigma) < |\sigma|$) can be seen to be linear-complete. Hence it is $rK$-complete. This is a result that was previously demonstrated through a direct argument by Zhenhao Li and the first author.

> Kolmogorov's set of nonrandom strings has maximum initial segment complexity amongst the c.e. sets.

Second (this observation was suggested by one of the referees), consider the set $S_\Omega$ containing the strings that lie lexicographically to the left of Chaitin's halting probability $\Omega$. This set is not $rK$-complete. Indeed, the segment $S\upharpoonright_{2^n}$ (including the codes for the strings of up to length n) has plain Kolmogorov complexity at most $C(\Omega\upharpoonright_{n+1})$. However there are infinitely many $n$ such that the latter is less than $n - \log(n)$ (this is a feature of every set, e.g. see [DH10, Corollary 3.11.3]). So by Theorem 2.7, the set $S_\Omega$ is not $rK$-complete.

## 3. Splitting theorems for computably enumerable sets

The decomposition of c.e. sets into disjoint c.e. sets is a major topic in computability theory (and in particular, the study of the lattice of the c.e. sets under inclusion). Various splitting theorems have been discovered since the early days of the subject and [DS93] is a comprehensive survey of this area. In this section we are concerned with splitting theorems that are relevant to the measures of complexity $\leq_{rK}, \leq_C, \leq_K$ and the results of Section 2. According to [Ste11, Chapter 2], [Bar11a, Section 5] (and the classic Sacks splitting theorem) if $\leq_r \in \{\leq_{rK}, \leq_C, \leq_K\}$ then every c.e. set $A >_r \emptyset$ is the disjoint union of two c.e. sets $A_0, A_1$ such that $A_i \not\leq_r A_{1-i}$ for each $i < 2$. On the other hand there are c.e. sets that cannot be split in the same degree, for example with respect to the Turing reducibility (a result from [Lac67]). We show that the same holds for $\leq_C$ (hence, also for $\leq_{rK}$ and $\leq_S$). If $\leq_r$ is a preorder, we let $\equiv_r$ denote the induced equivalence relation.

**Theorem 3.1** (Non-splitting for $\leq_C$)**.** *There exists a c.e. set $A$ such that for all pairs of c.e. sets $W, V$, if $W \cup V = A$ and $W \cap V = \emptyset$ then $A \not\equiv_C W$ or $A \not\equiv_C V$.*

**Proof.** In [Lac67] a version of this theorem was shown for $\equiv_C$ replaced by $\equiv_T$. In particular, there exists a c.e. set $S$ such that for all pairs of c.e. sets $W, V$, if $W \cup V = S$ and $W \cap V = \emptyset$ then $S \not\equiv_T W$ or $S \not\equiv_T V$. Let $A = \{2^{2^n} \mid n \in S\}$. It is straightforward to see that this computable translation of $S$ has the same properties: for all pairs of c.e. sets $W, V$, if $W \cup V = A$ and $W \cap V = \emptyset$ then $A \not\equiv_T W$ or $A \not\equiv_T V$. By [MS07] (as remarked in Section 1.1) for such a set $A \not\equiv_T W$ implies $A \not\equiv_C W$ and $A \not\equiv_T V$ implies $A \not\equiv_C V$. Hence $A$ has the desired properties. □

We note that the set of Theorem 3.1 may be chosen in the Turing degree of the halting problem, in the same way that the corresponding set from [Lac67] may be chosen with the same property (as noticed in [Lad73b]).

In view of the above discussion (in particular Theorem 3.1), it is rather surprising that every c.e. set can be split into two c.e. sets in the same $K$-degree. Moreover, as we see in the following, this unexpected fact has interesting consequences.

**Theorem 3.2** (Splitting in the same degree for $K$-reducibility)**.** *Every c.e. set can be split into two c.e. sets of the same $K$-degree. In other words, if $A$ is a c.e. set then there exist c.e. sets $A_0, A_1$ such that $A_0 \cup A_1 = A$, $A_0 \cap A_1 = \emptyset$ and $A \equiv_K A_0 \equiv_K A_1$.*

**Proof.** We fix a computable enumeration of $A$ and define the splitting $A_0$, $A_1$ as in the statement of the theorem. It suffices to construct a prefix-free machine $M$ such that the following requirements are met for all $n$:

$$K_M(A \restriction_n) \leq K(A_i \restriction_n). \tag{3.1}$$

Without loss of generality we may assume that enumerations into $A$ happen only at odd stages and that at each stage at most one such enumeration takes place. Also we may fix a universal prefix-free machine $U$, which is used for the definition of $K$-reducibility and which has weight less than $1/4$, and it is convenient to assume that new descriptions only appear in $U$ at even stages. At each odd stage $s+1$ we will be concerned with the weights

$$w_i(n)[s] = \sum_{n<k\leq s} 2^{-K(A_i\restriction_k)[s]}.$$

In the above, $K(A_i \restriction_k)[s]$ denotes the value as defined at the *end* of stage $s$.
*Construction.* At each odd stage $s+1$, if $n$ enters $A$ at this stage let $j$ be (the least number) such that $w_j(n)[s] \leq w_{1-j}(n)[s]$. Then enumerate $n$ into $A_{1-j}$. At each even stage $s+1$ and for each $n \leq s$ and $i \leq 1$ such that $K_M(A \restriction_n)[s] > K(A_i \restriction_n)[s]$ enumerate an $M$-description of $A[s] \restriction_n$ of length $K(A_i \restriction_n)[s]$.
*Verification.* By the construction, it suffices to show that the requests that we enumerate for $M$ have weight at most 1. Each request enumerated for $M$ at stage $s+1$ is triggered by finding $K_M(A \restriction_n)[s] > K(A_i \restriction_n)[s]$ for some $n$ and some $i = 0, 1$. In this way we may divide $M$ into two machines $M_0, M_1$ corresponding to $A_0, A_1$ respectively. We show that the weight of the $M_0$-requests is at most $1/2$. A symmetric argument shows that the weight of the $M_1$-requests is also at most $1/2$, so this will conclude the proof.

Each $M_0$-request at stage $s+1$ is triggered by finding $K_M(A \restriction_n)[s] > K(A_0 \restriction_n)[s]$ for some $n$, i.e. the request may be thought of as corresponding to the (leftmost) shortest $U[s]$-description $\tau$ of $A_0 \restriction_n [s]$. When this happens we say that $\tau$ becomes *used* (with respect to $M_0$). Once $\tau$ becomes used it remains used for the rest of the construction. Let us say that an $M_0$-request is *primary* if it corresponds to a $U$-description, which was not used prior to the point at which the request was made. If an $M_0$ request is not primary, we call it *secondary.* Note that secondary requests may occur when a number is enumerated into $A$ but not $A_0$, meaning that a new initial segment of $A$ must now be given a description as short as that which we have previously seen given for the unchanged initial segment of $A_0$. Clearly the weight of the primary $M_0$-requests is bounded by the weight of the universal machine (which determines $K(A_i \restriction_n)$ and its approximations). Since the latter is less than $1/4$, it suffices to show that the same holds for the weight of the secondary $M_0$-requests.

Note that if at an odd stage $s$ no number is enumerated into $A$ then any $M_0$-requests made at stage $s+1$ will be primary. Moreover the same holds if a number is enumerated in $A_0$ at stage $s$. We show that for every increase in the weight of the secondary $M_0$-requests we can count an equal (or even larger) increase in the weight of the universal machine $U$. Indeed, if at stage $s+1$ some secondary $M_0$-requests are enumerated, a number $m$, which is smaller than all the lengths of the strings for which these secondary requests require descriptions, must have entered $A_1$ at stage $s$. According to the construction (and since we assume that new descriptions only enter $U$ at even stages) this means that $w_0(m)[s] = w_0(m)[s-1] \leq w_1(m)[s-1]$. Hence we can count weight $w_1(m)[s-1]$ in the domain of $U$, which is greater than or equal to the total weight of the secondary $M_0$-requests made at stage $s+1$. Since $A(m)[t] \neq A(m)[s-1]$ for all $t \geq s$ this weight in the domain of $U$ will not be counted twice. It follows that the weight of the secondary $M_0$-requests is also bounded by $1/4$. Hence the weight of the $M_0$-requests is bounded by $1/4+1/4=1/2$. This (and the entirely symmetric argument for $M_1$) shows that the weight of the $M$-requests is bounded by 1. $\square$

Clearly if a set is complex then all sets that are $K$-equivalent to it are complex. Hence the following is a consequence of Theorem 3.2.

**Corollary 3.3** (Splitting for complex c.e. sets)**.** *Every c.e. complex set can be split into two disjoint complex c.e. sets.*

The complex c.e. sets are exactly the c.e. sets in the weak truth table degree of the halting problem. Hence Corollary 3.3 may be stated as follows.

**Corollary 3.4** (Splitting for wtt-complete sets)**.** *Every c.e. set $A \equiv_{wtt} \emptyset'$ is the disjoint union of two c.e. sets $A_0 \equiv_{wtt} A_1 \equiv_{wtt} \emptyset'$.*

We note that Corollary 3.4 is no longer true if $\equiv_{\mathrm{wtt}}$ is replaced by $\equiv_{\mathrm{T}}$. Moreover, there are many c.e. sets $X$ such that Corollary 3.4 is no longer true when $\emptyset'$ is replaced by $X$ (even when $\equiv_{\mathrm{wtt}}$ is replaced with $\equiv_{\mathrm{T}}$). These are results from [Lac67, Lad73b]. The existence of a c.e. Turing degree such that all of its c.e. members can be split in the same degree is a result from [Lad73a]. Such degrees are known as 'completely mitotic' (*mitotic* being the name of c.e. sets that can be split in the same Turing degree). Since the Turing degree of the halting problem is not completely mitotic, Corollary 3.4 is an interesting consequence. Further contrasts may be obtained by comparing our corollary with a number of splitting results that are surveyed in [DS93]. A particularly striking example (originally from [AS85]) is the existence of a Turing complete set such that every c.e. splitting of it has a member which is low.

We conclude with a generalization of the splitting theorem from [Ste11, Chapter 2], [Bar11a, Section 5] that was discussed at the beginning of this section. We present two versions. The first version is as follows.

**Theorem 3.5** (Splitting for the $K$-degrees and the $C$-degrees, version 1)**.** *Let $f$ be a right c.e. function and let $A$ be a c.e. set such that $\neg\exists c \forall n,\ C(A\restriction_n) \le f(n)+c$. Then $A$ is the disjoint union of two c.e. sets $A_0, A_1$ such that $\neg\exists c\forall n,\ C(A_i\restriction_n) \le f(n)+c$ for each $i < 2$. A corresponding statement holds for $K$ in place of $C$.*

**Proof.** By Theorem 3.2 the $K$-version of the above theorem holds. For the $C$-version we argue as follows. Let $f[s]$ be a computable approximation to $f$ such that $f(n)[s+1] \le f(n)[s]$ for all $n, s$. In the course of enumerating the elements of $A$ into $A_0$ and $A_1$ we satisfy the following requirement for $e \in \omega$ and $i < 2$.

$$R_{\langle e,i\rangle} : \exists n\ \big[C(A_i\restriction_n) > f(n) + e\big].$$

Define the *length of agreement* of $R_{\langle e,i\rangle}$ at stage $s$ by

$$l(e,i)[s] = \text{the largest } n \le s \text{ such that } \forall j \le n\ (C(A_i\restriction_j)[s] \le f(j)[s] + e)$$

and let $r(e,i)[s] = \max_{t\le s}\{l(e,i)[t], e\}$. By definition, $r(e,i)[s]$ is non-decreasing in the stages $s$. Let $A_i[0] = \emptyset$ for $i = 0, 1$ and without loss of generality assume that at each stage exactly one element is enumerated in $A$.

*Construction.* If $x \in A[s+1] - A[s]$ consider the least $\langle e,i\rangle$ such that $x \le r(e,i)[s]$ and enumerate $x$ into $A_i$.

*Verification.* Clearly $A_0, A_1$ are c.e. and disjoint; moreover $A = A_0 \cup A_1$. By induction we show that each $R_{\langle e,i\rangle}$ is satisfied. Clearly $R_{\langle e,i\rangle}$ is met if and only if $r(e,i)[s]$ reaches a limit as $s \to \infty$. The induction hypothesis is that there exists some stage $s_0$ such that for all $\langle e',i'\rangle < \langle e,i\rangle$ requirement $R_{\langle e',i'\rangle}$ is met and $r(e',i')[s]$ remains constant for all $s \geq s_0$; moreover all numbers enumerated in $A$ after $s_0$ are larger than the final values of $r(e',i')$, $\langle e',i'\rangle < \langle e,i\rangle$.

For a contradiction, suppose that $R_{\langle e,i\rangle}$ is not met, so $r(e,i)[s]$ tends monotonically to infinity as $s \to \infty$. Then by the construction, $A_{1-i}$ is computable. Hence $\exists c \forall n,\ C(A\upharpoonright_n) \leq C(A_i \upharpoonright_n) + c$. This relation, along with the failure of $R_{\langle e,i\rangle}$ contradicts the hypothesis about $A$. Hence $R_{\langle e,i\rangle}$ is met and this concludes the induction step. □

A class of sets of very low initial segment complexity was introduced in [LL99] by the name *ultracompressible sets*. A set $A$ is ultracompressible if for every computable order $f$ there exists $c$ such that $\forall n,\ K(A\upharpoonright_n) \leq K(n) + f(n) + c$. The following is an immediate consequence of Theorem 3.2.

**Corollary 3.6.** *Every c.e. set that is not ultracompressible is the disjoint union of two c.e. sets that are not ultracompressible.*

A similar argument gives the following version of Theorem 3.5.

**Theorem 3.7** (Splitting for the $K$-degrees and the $C$-degrees, version 2)**.** *Let $f$ be a right c.e. function and let $A$ be a c.e. set such that $\neg\exists c \forall n,\ K(A\upharpoonright_n) \leq K(n) + f(n) + c$. Then $A$ is the disjoint union of two c.e. sets $A_0, A_1$ such that $\neg\exists c\forall n,\ K(A_i\upharpoonright_n) \leq K(A_{1-i}\upharpoonright_n) + f(n) + c$ for each $i < 2$. Moreover a corresponding statement holds for $C$ in place of $K$.*

**Proof.** We give the proof for $K$-reducibility, since the case for $C$-reducibility is entirely similar. We proceed as in the proof of Theorem 3.5 (using the same notation).

$$R_{\langle e,i\rangle} : \exists n\ \big[K(A_{1-i}\upharpoonright_n) > K(A_i\upharpoonright_n) + f(n) + e\big].$$

Define the *length of agreement* $l(e,i)[s]$ of $R_{\langle e,i\rangle}$ at stage $s$ to be the largest $n \leq s$ such that $\forall j \leq n\ (K(A_{1-i}\upharpoonright_j)[s] \leq K(A_i\upharpoonright_j)[s] + f(j)[s] + e)$. Define the *restraint* $r(e,i)[s]$ of $R_{\langle e,i\rangle}$ at stage $s$ to be $\max_{t\leq s}\{l(e,i)[t], e\}$. By definition, $r(e,i)[s]$ is non-decreasing in the stages $s$. Let $A_i[0] = \emptyset$ for $i = 0,1$ and without loss of generality assume that at each stage exactly one element is enumerated in $A$.

*Construction.* If $x \in A[s+1] - A[s]$ consider the least $\langle e,i\rangle$ such that $x \leq r(e,i)[s]$ and enumerate $x$ into $A_{1-i}$.

*Verification.* Clearly $A_0, A_1$ are c.e. and disjoint; moreover $A = A_0 \cup A_1$. By induction we show that each $R_{\langle e,i\rangle}$ is satisfied. Clearly $R_{\langle e,i\rangle}$ is met if and only if $r(e,i)[s]$ reaches a limit as $s \to \infty$. The induction hypothesis is that there is a stage $s_0$ such that for all $\langle e',i'\rangle < \langle e,i\rangle$ requirement $R_{\langle e',i'\rangle}$ is met and $r(e',i')[s]$ remains constant for all $s \geq s_0$; moreover all numbers enumerated in $A$ after $s_0$ are larger than the final values of $r(e',i')$, $\langle e',i'\rangle < \langle e,i\rangle$.

If $R_{\langle e,i\rangle}$ is not met then $A_i$ is computable. Hence there exists some constant such that $K(A\upharpoonright_n) \leq K(n) + f(n) + c$. This contradicts the hypothesis on $A$. Hence $R_{\langle e,i\rangle}$ is met and the induction step is complete. □

We conclude with a couple of open problems. The following question from [RS08] is related to the topics discussed in this note and remains open.

**Question 1.** *Is every sequence $rK$-reducible to a random sequence?*

Another question is whether, for example, the $K$-degrees of the c.e. sets are dense. Given the coding method that we introduced in the proof of Theorem 2.2, this seems to be related to the question of whether every pair of c.e. sets has a least upper bound in the $K$-degrees. We note that the density results that were obtained in [DHN02] are not very related to this question as they refer to the case of c.e. reals (where it is easy to see that the usual addition between reals is a join operator) and not c.e. sets.

## References


[AS85] Klaus Ambos-Spies. Anti-mitotic recursively enumerable sets. *Mathematical Logic Quarterly*, 31(29-30):461–477, 1985.

[Bar68] Janis M. Barzdin. Complexity of programs to determine whether natural numbers not greater than n belong to recursively enumerable set. *Soviet Math. Dokl.*, 9:1251–1254, 1968.

[Bar05] George Barmpalias. Computably enumerable sets in the Solovay and the strong weak truth table degrees. In S. Barry Cooper, Benedikt Löwe, and Leen Torenvliet, editors, *CiE*, volume 3526 of *Lecture Notes in Computer Science*, pages 8–17. Springer, 2005.

[Bar11a] George Barmpalias. On strings with trivial Kolmogorov complexity. *Int J Software Informatics*, 5(4):609–623, 2011.

[Bar11b] George Barmpalias. Universal computably enumerable sets and initial segment prefix-free complexity. Submitted, 2011.

[BL11] George Barmpalias and Angsheng Li. Kolmogorov complexity and computably enumerable sets. Submitted, 2011.

[Cha75] Gregory J. Chaitin. A theory of program size formally identical to information theory. *J. Assoc. Comput. Mach.*, 22:329–340, 1975.

[CHKW01] Cristian Calude, Peter Hertling, Bakhadyr Khoussainov, and Yongge Wang. Recursively enumerable reals and Chaitin Ω numbers. *Theoret. Comput. Sci.*, 255(1-2):125–149, 2001.

[DH10] Rod Downey and Denis Hirschfeldt. *Algorithmic Randomness and Complexity.* Springer, 2010.

[DHL04] Rod G. Downey, Denis R. Hirschfeldt, and Geoff LaForte. Randomness and reducibility. *J. Comput. System Sci.*, 68(1):96–114, 2004.

[DHN02] Rod Downey, Denis Hirschfeldt, and André Nies. Randomness, computability and density. *SIAM J. Computing*, 31:1169–1183, 2002.

[DS93] Rod Downey and Michael Stob. Splitting theorems in recursion theory. *Ann. Pure Appl. Logic*, 65(1):106 pp, 1993.

[Kan69] M. I. Kanovich. On the decision complexity of algorithms. *Soviet Mathematics Doklady*, 10:700–701, 1969. And 11:17–20, 1970.

[KHMS11] Bjørn Kjos-Hanssen, Wolfgang Merkle, and Frank Stephan. Kolmogorov complexity and the recursion theorem. *Trans. Amer. Math. Soc.*, 363, 2011.

[KS01] Antonín Kučera and Theodore Slaman. Randomness and recursive enumerability. *SIAM J. Comput.*, 31(1):199–211, 2001.

[Kum96] Martin Kummer. Kolmogorov complexity and instance complexity of recursively enumerable sets. *SIAM J. Comput.*, 25(6):1123–1143, 1996.

[Lac67] Alistair H. Lachlan. The priority method I. *Z. Math. Logik Grundlag. Math.*, 13:1–10, 1967.

[Lad73a] Richard Ladner. A completely mitotic nonrecursive r.e. degree. *Trans. Amer. Math. Soc.*, 184:497–507, 1973.

[Lad73b] Richard Ladner. Mitotic recursively enumerable sets. *J. Symbolic Logic*, 38:199–211, 1973.

[Lev73] Leonid A. Levin. The concept of a random sequence. *Dokl. Akad. Nauk SSSR*, 212:548–550, 1973.

[LL99] James I. Lathrop and Jack H. Lutz. Recursive computational depth. *Inf. Comput.*, 153(1):139–172, 1999.

[MS07] Wolfgang Merkle and Frank Stephan. On C-degrees, H-degrees and T-degrees. In *Twenty-Second Annual IEEE Conference on Computational Complexity (CCC 2007), San Diego, USA, 12–16 June 2007*, pages 60–69, Los Alamitos, CA, USA, 2007. IEEE Computer Society.

[Nie09] André Nies. *Computability and Randomness*. Oxford University Press, 2009.

[Odi89] P. G. Odifreddi. *Classical recursion theory. Vol. I*. North-Holland Publishing Co., Amsterdam, 1989.

[RS08] Alexander Raichev and Frank Stephan. A minimal *rk*-degree. In *Lecture Notes Series Vol. 15, Institute for Mathematical Sciences, National University of Singapore*, pages 261–269, 2008.

[Sol75] Robert Solovay. Handwritten manuscript related to Chaitin's work. IBM Thomas J. Watson Research Center, Yorktown Heights, NY, 215 pages, 1975.

[Ste11] Tom Sterkenburg. *Sequences with trivial initial segment complexity*. MSc Dissertation, University of Amsterdam, February 2011.

**George Barmpalias:** State Key Lab of Computer Science, Institute of Software, Chinese Academy of Sciences, Beijing 100190, P.O. Box 8718, China.
*E-mail address*: barmpalias@gmail.com
*URL*: http://www.barmpalias.net

**Rupert Hölzl**, Bureau 6A35 LIAFA, 175 rue du Chevaleret, 75013 Paris, France.
*E-mail address*: r@hoelzl.fr
*URL*: http://hoelzl.fr

**Andrew E.M. Lewis**, School of Mathematics, University of Leeds, LS2 9JT Leeds, United Kingdom.
*E-mail address*: andy@aemlewis.co.uk
*URL*: http://aemlewis.co.uk

**Wolfgang Merkle**, Institut für Informatik, Ruprecht-Karls-Universität, Heidelberg, Germany.
*E-mail address*: merkle@math.uni-heidelberg.de
*URL*: http://www.math.uni-heidelberg.de/logic/merkle/merkle.html