The Complexity of the Set of Nonrandom Numbers

Frank Stephan
Department of Mathematics and School of Computing
National University of Singapore, 2 Science Drive 2, Singapore 117543
Email: fstephan@comp.nus.edu.sg

Let C and H denote the plain and prefix-free Kolmogorov complexity, respectively. Then the sets NRC of nonrandom numbers with respect to C has neither a maximal nor an r-maximal superset. The set of NRH of nonrandom numbers with respect to H has an r-maximal but no maximal superset. Thus the lattices of recursively enumerable supersets (modulo finite sets) of NRC and NRH are not isomorphic. Further investigations deal with the related set NRW of numbers x with a stronger nonrandomness property: $x \neq \max\{W_e\}$ for any $e < x$ where $W_0, W_1, \ldots$ is derived from the underlying acceptable numbering of partial-recursive functions. Friedman originally asked whether $NRW \equiv_T K$ for every underlying acceptable numbering and Davie provided a positive answer for many underlying acceptable numberings. Later Teutsch asked whether the set NRW can be r.e. or co-r.e.; as an answer to this question it is shown that in the case that the underlying numbering is a Kolmogorov numbering, NRW is not n-r.e. for any n. If one uses any acceptable numbering instead of a Kolmogorov numbering, then the underlying numbering can be chosen such that NRW is a co-2-r.e. set; but it cannot be a 2-r.e. set for any acceptable numbering.

1. Introduction

Let C and H denote the plain and prefix-free Kolmogorov complexity, respectively. Furthermore, one can identify the numbers in the set $I_n = \{2^n - 1, 2^n, 2^n + 1, \ldots, 2^{n+1} - 2\}$ with the binary strings of length n; a number x corresponds to a string σ iff $x + 1$ is the binary value of the string 1σ. For having an easier connection to other fields of recursion theory, natural numbers are used from now on. For any number x, the n with $x \in I_n$ is called the length of x, written $|x|$. The plain Kolmogorov complexity C is defined as

$$C(x) = \min\{|p| : U(p) = x\}$$

where U is a universal function. That is, the value C based on U satisfies the following two conditions:

(1) The range of U is the set of all natural numbers, so that $C(x)$ is defined for all x.
(2) For every further unary partial-recursive function V there is a constant c with $C(V(p)) \leq |p| + c$ for all p in the domain of V.

The other variant H is based on prefix-free machines. Chaitin [1, 2, 3] and Levin [11] laid the foundations and showed the significance of this alternative approach which developed together with the original notion C to the two best accepted concepts in Kolmogorov complexity. A prefix-free machine satisfies that all different strings p, q in its domain are incomparable with respect to the string-prefix-relation; alternatively, one can also use the Kraft-Chaitin-Theorem [1, 2, 3, 11] and say that a machine U is equivalent to a prefix-free one iff

$$\sum_{p \in dom(U)} 2^{-|p|} \leq 1.$$

A prefix-free machine U is universal iff the two conditions above hold where in Condition 2 only prefix-free machines V are considered. Then

$$H(x) = \min\{|p| : U(p) = x\}$$

where U is a prefix-free partial-recursive unary function which is universal for the class of all prefix-free partial-recursive unary functions. Now the two sets in question are defined as

- $NRC = \{x : C(x) < |x|\}$;
- $NRH = \{x : H(x) < |x|\}$.

They are called the sets of nonrandom numbers with respect to C and H or the sets of compressible strings with respect to C and H. These sets are standard examples of simple sets, which do not arise through a complicated construction but are just given as natural. Furthermore, they are wtt-complete but not btt-complete. Kummer [9] showed that NRC is tt-complete (for any given universal machine), but according to Muchnik and Peretselski [14] the tt-completeness of *NRH* depends on the choice of the universal machine. The existence of universal machines for which *NRH* is tt-complete is easy to show; a clever definition makes it possible to satisfy the equivalence

$$n \in K \Leftrightarrow |NRH \cap \{2^n, 2^n+1, \ldots, 2^{n+1}-1\}| \text{ is odd.}$$

The choice of a universal machine where *NRH* is not tt-complete is the more difficult part.

In the present work, it is shown that the sets *NRC* and *NRH* can also be used as examples for the following theorems. Martin [13] showed that there is a recursively enumerable coinfinite set without a maximal supersets; both sets *NRC* and *NRH* have this property as well. Lachlan [10] and Robinson [18] constructed a recursively enumerable coinfinite set without an r-maximal superset; *NRC* is also such an example. Furthermore, Lachlan and Robinson constructed a set with an r-maximal but without any maximal superset; *NRH* is such an example.

For these theorems, recall that a recursively enumerable and coinfinite set A is called maximal if either $\overline{A} \subseteq^* B$ or $\overline{A} \subseteq^* \overline{B}$ for every recursively enumerable set B. Furthermore, a recursively enumerable and coinfinite set A is called r-maximal if either $\overline{A} \subseteq^* B$ or $\overline{A} \subseteq^* \overline{B}$ for every recursive set B. Here $X \subseteq^* Y$ if $X - Y$ is finite, that is, if almost all $x \in X$ are also in Y; furthermore, $X \subset^* Y$ iff $X \subseteq^* Y$ and $Y \not\subseteq^* X$.

A lot of variants of the basic notions C and H have been studied in the theory of Kolmogorov complexity. One related notion to the sets *NRC* and *NRH* is the set

$$NRW = \{x : \exists e < x \; [x = \max(W_e \cup \{0\})]\}$$

which is based on the underlying acceptable numbering of partial-recursive functions with W_e being the domain of the e-th partial-recursive function. $W_0, W_1, \ldots$; Friedman [6] asked what the Turing-degree of *NRW* is; Davie [4] obtained the by now best results with respect to this problem. Davie first solved this question for the case where the underlying numbering is a Kolmogorov numbering. Here a numbering $\varphi_0, \varphi_1, \ldots$ of partial-recursive functions is a Kolmogorov numbering iff, given any other numbering $\psi_0, \psi_1, \ldots$ of partial-recursive functions, there are constants c, d such that one can compute for every index n an index $m \leq cn + d$ with $\varphi_m = \psi_n$. Later, based on correspondence with Solovay, Davie pointed out how to solve Friedman's question for a further large class of acceptable numberings. But Friedman's question is still not completely solved. At the end, an alternative proof using Kummer's Cardinality Theorem is given for Davie's first result that $K \leq_T NRW$ whenever the underlying numbering is a Kolmogorov numbering.

Schaefer and Teutsch also studied related problems, for example, whether *NRW* is an r.e. or a co-r.e. set [20]. Note that the set *NRW* is similar to *NRC* and *NRH*; so if one chooses the universal machines and the numbering $W_0, W_1, \ldots$ properly, then one gets $NRH \subseteq NRC \subseteq NRW$ which would directly imply that *NRW* is co-immune and thus not a co-r.e. set. But the obtained results are more general: There is an acceptable numbering such that — when based on this numbering — *NRW* is a co-2-r.e. set. But *NRW* cannot be a 2-r.e. set. Furthermore, if *NRW* is based on a Kolmogorov numbering then *NRW* is not n-r.e. for any number n. In any case, *NRW* is still ω-r.e. as witnessed by the approximation A_s with $x \in A_s$ iff there is an $y < x$ with $\{x\} \subseteq W_{y,s} \cap \{0, 1, \ldots, s\} \subseteq \{0, 1, \ldots, x\}$.

Unexplained recursion-theoretic notation follows the books of Odifreddi [15, 16] and Soare [19].

2. Known Results

Neither the set *NRC* nor the set *NRH* have hyperhypersimple supersets. Defining the universal machines adequately gives that $C(x) \leq H(x)$ for all x and thus $NRH \subseteq NRC$; this will be assumed here and in all further proofs.

Remark 2.1: Martin [13] proved that there is a recursively enumerable coinfinite set without a maximal superset. Odifreddi [16] proved Proposition IX.2.27 by producing a set A which contains for all n and all $e \leq n$ every element of a set $I_n - W_e$ whenever that set has at most one element y. He then showed that such a set A has no maximal superset. As

$$C(y) \leq C((n+e)(n+e+1)+n) + c \leq 2\log(n+1)$$

for some constant c independent of n, e whenever such a y exists, the construction implies that for every function f with $f(n) \geq 3 \cdot \log(n)$ for almost all n, the sets

$$\{x : C(x) < f(x)\} \text{ and } \{x : H(x) < f(x)\}$$

do not have a maximal superset. In particular *NRC* and *NRH* do not have maximal supersets. One can generalize the above result to all recursive, increasing and unbounded functions f by using larger intervals J_n instead of I_n. Taking $J_n = \{x : f(x) \in I_n\}$, one gets that the complement of every maximal set contains infinitely many x with $C(x) < f(x)$.

So this result is already more or less there. Nevertheless, a formal proof will be given for the sake of completeness. The Kolmogorov complexity part of this proof can be put into the following form, which is quite often used implicitly although the author is not aware of a direct reference.

Proposition 2.2: *Let A be a recursively enumerable set. Then there is a constant c such that the following holds:*
If $NRC \subseteq A$ then for all n, either $I_n \subseteq A$ or $|I_n - A| \geq 2^{-c} \cdot |I_n|$.
If $NRH \subseteq A$ then for all n, either $I_n \subseteq A$ or $|I_n - A| \geq 2^{-2\cdot|n|-c} \cdot |I_n|$.

Proof: Assume that A is recursively enumerable and $NRC \subseteq A$.

Then one splits each interval I_n into intervals $J_{n,m}$ with 2^{n-m} elements for $m = 1, 2, \ldots, n$ plus one further element. Now one can define a partial-recursive function V such that $V[J_{n,m}] = \{V(p) : p \in J_{n,m}\} \supseteq I_{n+m} - A$ whenever $|I_{n+m} - A| \leq 2^{n-m}$ by waiting on each interval $J_{n,m}$ for the event $|I_{n+m} - A| \leq |J_{n,m}|$ to take place and by then assigning the values correspondingly. Note that V is not total as there are intervals $J_{n,m}$ where the corresponding event will never take place. It follows that $C(x) \leq n - m$ for all $x \in I_{n+m}$ whenever $|I_{n+m} - A| \leq 2^{n-m}$. Now one can argue that there is a constant c' with $C(V(p)) < |p| + c'$ for all p in the domain of V. It follows that $I_{n+c'}$ is not in the range of $V[J_{n,c}]$ and hence there is no n such that $1 \leq |I_{n+c'} - A| \leq 2^{n-c'}$.

The second result is just based on the fact that $|H(x) - C(x)| \leq 2\cdot|n| + c''$ for some constant c'', all n and all $x \in I_n$ [12]. So the whole theorem holds with $c = 2(c' + c'')$. □

Proposition 2.3: *Neither the set NRC nor the set NRH have a hyper-hypersimple superset.*

Proof: As $NRH \subseteq NRC$, it is sufficient to show that no coinfinite r.e. superset A of NRH is hyperhypersimple. So let such a coinfinite r.e. superset A of NRH be given. By Proposition 2.2 there is an increasing and unbounded recursive function b such that for infinitely many n the set $I_n - A$ contains at least $b(n)$ many elements.

Having this property, it is easy to define a partial-recursive function ψ such that ψ takes on every set I_n with $|I_n - A| \geq b(n)$ on $b(n)$ elements of $I_n - A$ the values $0, 1, \ldots, b(n) - 1$, respectively. Whenever then some

$x \in I_n$ with $\psi(x) = m \wedge m < b(n)$ is enumerated into A_{s+1} at stage s, one takes the least value $x \in I_n - A_{s+1}$ where $\psi_s(x)$ is still undefined and assigns the value $\psi_{s+1}(x) = m$.

By choice of the function b there are for every m infinitely many n such that $|I_n - A| \geq n > m$ and $\psi(x) = m$ for an $x \in I_n - A$. It follows from a result of Yates [21] that A is not hyperhypersimple. □

The next result is an alternative proof for the result of Robinson [18] in 1967 and Lachlan [10] in 1968 that there is a recursively enumerable and coinfinite set without an r-maximal superset. The proof is based on an adjustment of Lachlan's proof [10].

Theorem 2.4: *NRC has no r-maximal superset.*

Proof: Let A be a recursively enumerable and coinfinite superset of *NRC*. By Proposition 2.2 there is a constant c such that, for infinitely many n, $I_n - A$ has at least 2^{n-c} elements. Now let

$$B_m = \{m, m + 2^{c+1}, m + 2 \cdot 2^{c+1}, m + 3 \cdot 2^{c+1}, \ldots\}$$

for $m = 0, 1, \ldots, 2^{c+1} - 1$. These sets form a finite partition of the natural numbers, but for each m, $\overline{A} \not\subseteq^* B_m$ as

$$\exists^\infty n \,[\, |I_n - A| > 2^{-c-1} \cdot |I_n| \geq |B_m \cap I_n| \,].$$

Hence A is not r-maximal. □

3. *NRH* has r-maximal supersets

One main result of the present work is that in contrast to *NRC*, *NRH* has an r-maximal superset.

Theorem 3.1: *There is an r-maximal superset of NRH.*

Proof: An r-maximal set A is constructed such that $A \cup NRH$ is coinfinite. Then $A \cup NRH$ is also r-maximal and a superset of *NRH*.

Now let $J_0, J_1, J_2, \ldots$ be a partition of the natural numbers such that each J_n has 2^{3n} elements. Let $L_n = \cup\{I_m : m \in J_n\}$ be the union of all I_m such that $m \in J_n$. The sets $L_0, L_1, L_2, \ldots$ are also a recursive partition of the natural numbers.

First, a recursively enumerable set B is defined such that for all r.e. sets W_i, W_j and $n > i + j$, if $L_n - B \subseteq W_i \cup W_j$ then either $L_n - B \subseteq W_i$ or $L_n - B \subseteq W_j$. This is obtained by making one indirect step into Lachlan's construction. More precisely, $L_n - B$ is forced to become a subset of W_e whenever W_e contains at stage s the majority of the sets I_k with $I_k \not\subseteq B_s$ satisfy that at least the half of $I_k - B_s$ is contained in $W_{e,s}$. More formally, the following is done.

- For each n and stage s, let e be the remainder of s divided by n and do the following:
- Let $D_0 = \{k \in J_n : I_k \not\subseteq B_s\}$;
- Let $D_1 = \{k \in D_0 : |(I_k - B_s) \cap W_{e,s}| \geq 0.5 \cdot |I_k - B_s|\}$;
- If $2|D_1| \geq |D_0|$ and $L_n - B_s \not\subseteq W_{e,s}$ then update B_s on L_n by letting
    - $B_{s+1} \cap I_k = I_k$ for all $k \in D_0 - D_1$;
    - $B_{s+1} \cap I_k = (B_s \cap I_k) \cup (I_k - W_{e,s})$ for all $k \in D_1$;

  else let $B_{s+1} \cap L_n = B_s \cap L_n$.

Now one can verify that the desired property holds. Assume that $L_n \not\subseteq B$ but $L_n \subseteq W_i \cup W_j$ and t is so large that $B_t \cap L_n = B \cap L_n$, $W_{i,t} \cap L_n = W_i \cap L_n$ and $W_{j,t} \cap L_n = W_j \cap L_n$. In every stage $s \geq t$, $D_0 = \{k \in J_n : I_k \not\subseteq B\}$ as $B_s \cap L_n = B \cap L_n$. Furthermore,

$$\forall k \in D_0 \; \exists e \in \{i, j\} \; [\, |(I_k - B_s) \cap W_{e,s}| \geq 0.5 \cdot |I_k - B_s| \,].$$

Thus, either for $e = i$ or for $e = j$, it must hold that at stage $s = tn + e$ that $2|D_1| \geq |D_0|$. As at this step, no new elements are put into B, the condition $L_n - B_s \subseteq W_{e,s}$ must be satisfied. So $L_n - B \subseteq W_i$ or $L_n - B \subseteq W_j$.

Second, it is shown that for almost every n, $L_n \not\subseteq B$. To see this, note that for each interval L_n it happens only in at most n stages s that some elements of L_n are enumerated into B. One can easily verify that after each step, at least half of the $k \in J_n$ with $I_k \not\subseteq B_s$ also satisfy the condition $|I_k B_{s+1}| \geq |I_k - B_s| \cdot 0.5$ while all other $k \in J_n$ satisfy $I_k \subseteq B_{s+1}$. As a consequence, in the limit, $|J_n| \cdot 2^{-n}$ of the intervals I_k with $k \in J_n$ satisfy $|I_k - B| \geq |I_k| \cdot 2^{-n} \geq 2^{k-n}$. As $|J_n| = 2^{3n}$, the number of these intervals is at least 2^{2n}.

Chaitin [2] showed that there is a constant c independent of k and r with $H(x) \leq k + H(k) - r$ for at most 2^{k+c-r} numbers $x \in I_k$; Downey and Hirschfeldt [5] call it the "Counting Theorem". Taking $r = n + c$ gives that at least $2^k - 2^{k-n}$ members x of I_k satisfy $H(x) > k + H(k) - n - c$ and thus $H(x) > k + n - c$. So, for sufficiently large n, the set J_n contains a k such that $H(k) \geq 2n$ and $I_k \not\subseteq B \cup \mathit{NRH}$. Hence $L_n \not\subseteq B \cup \mathit{NRH}$ for all sufficiently large n.

Third, taking any maximal set M and $A = B \cup \{x : \exists n \in M\, (x \in L_n)\}$, the sets A and $A \cup \mathit{NRH}$ are r-maximal. Clearly, A and $A \cup \mathit{NRH}$ are coinfinite by the preceding two paragraphs. For each e define the r.e. set

$$V_e = \{n : L_n - B \subseteq W_e\}.$$

Now let i, j be indices of r.e. sets such that W_j is the complement of W_i. By the first part of the proof, all numbers $n > i + j$ are in $V_i \cup V_j$. As M is maximal, $\overline{M} \subseteq^* V_e$ for some $e \in \{i, j\}$. Then $\overline{A} \subseteq^* W_e$. Therefore, A and $A \cup NRH$ are r-maximal. □

The r-maximal set $A \cup NRH$ constructed has by Theorem 2.3 no hyperhypersimple superset. Thus it has no maximal superset. Such type of r-maximal sets had been constructed by Robinson [18] in 1967 and by Lachlan [10] in 1968.

Corollary 3.2: *There is an r-maximal set without a maximal superset.*

A maximal set A has the property that the structure $\mathcal{L}_A^*$ of its r.e. superset modulo finite sets is just the 2-element Boolean Algebra. Hyperhypersimple sets are characterised as those sets where this structure is a Boolean Algebra. Theorems 2.4 and 3.1 show that the two sets of nonrandom numbers have a different superset structure.

Corollary 3.3: *Given a set E, let $(\mathcal{L}_E^*, \subset^*)$ be the partially ordered set of the r.e. supersets of E modulo finite sets with the ordering induced from set-inclusion. Then the structures $(\mathcal{L}_{NRC}^*, \subset^*)$ and $(\mathcal{L}_{NRH}^*, \subset^*)$ are not isomorphic.*

4. The Problems of Friedman and Teutsch

A numbering φ of partial recursive functions is called a Kolmogorov numbering iff for every further numbering of partial-recursive functions ψ there are constants c, d such that, for all x, every ψ_x equals some φ_y with $y < cx + d$. This translated of course to the domains W_x of φ_x: for every further numbering $A_0, A_1, A_2, \ldots$ of any r.e. sets there are constants c, d such that for every x there is an $y \leq cx + d$ with $W_y = A_x$. Furthermore, if $W_0, W_1, W_2, \ldots$ is derived from an acceptable numbering φ of partial recursive functions, then there is for every numbering $A_0, A_1, A_2, \ldots$ of any r.e. sets a recursive function f such that $W_{f(x)} = A_x$ for all indices x. Let NRW denote the set

$$NRW = \{x : \exists y < x\ [x = \max(W_y \cup \{0\})]\}$$

for a given fixed Kolmogorov numbering φ of the partial recursive functions. Friedman [6] asked whether $NRW \equiv_T K$ for every underlying acceptable numbering φ; certainly it is easy to construct some for which is true. Teutsch [20] added the question whether NRW can be r.e. or co-r.e. for some acceptable numbering. The general case is indeed the more difficult one and therefore only Teutsch's question can be answered for all acceptable numberings.

In the following, a set is 2-r.e. iff it is the difference of two r.e. sets, it is 3-r.e. iff it is the difference of an r.e. set minus a 2-r.e. set and so on. Alternatively, one can say that a set A is n-r.e. iff there is an approximation A_s of A such that $A_0 = \emptyset$ and there are for every n at most n indices with s with $A_{s+1}(x) \neq A_s(x)$. Recall the default approximation $A_0, A_1, A_2, \ldots$ is that

$$x \in A_s \Leftrightarrow \exists y < x \; [\{x\} \subseteq W_{y,s} \cap \{0, 1, \ldots, s\} \subseteq \{0, 1, \ldots, x\}].$$

This approximation makes for each x up to $2x$ mind changes; so it witnesses that NRW is ω-r.e. but it does not witness that NRW is n-r.e. for any natural number n. So the main question is whether NRW can be n-r.e. for some n. The answer is affirmative if the underlying numbering is not requested to be a Kolmogorov numbering.

Theorem 4.1: *There is an acceptable numbering such that NRW is the complement of some 2-r.e. set.*

Proof: Fix the recursive partition J_0, $\{x_0\}$, J_1, $\{x_1\}$, J_2, $\{x_2\}$ of the natural numbers satisfying $|J_k| = 2^k + 2$ for all k; this is the partition given by $J_0 = \{0, 1, 2\}$, $x_0 = 3$, $J_1 = \{4, 5, 6, 7\}$, $x_1 = 8$, $J_2 = \{9, 10, 11, 12, 13, 14\}$ and so on. Furthermore, let for every k the number

$$y_k = \min\{x \in J_k : \forall z \in J_k \; [z \leq x \vee H(z) < k]\}.$$

Note that the intervals J_k are chosen so large that every interval J_k contains at least three numbers z with $H(z) \geq k$ and y_k is just the maximum of all these z. Furthermore, the y_k are uniformly approximable from above as the formula of their definition shows, let $y_{k,s}$ be the value of y_k before stage s. Now define $A_s(x) = 1$ iff there is a k with $x \in J_k$ and $s > 0$ and $x = y_{k,s}$. It is easy to verify that this approximation witnesses that A is 2-r.e.: At the beginning, $A_s(x) = 0$; if $y_{k,s}$ has come down and reached x then $A_s(x) = 1$; if $y_{k,s}$ has gone below x then $A_s(x) = 0$ again.

The proof is completed by defining an acceptable enumeration $W_0, W_1, \ldots$ such that NRW based on this numbering is a finite variant of $\overline{A}$. So let any acceptable numbering ψ be given. Define W_{x_k} to be the domain of ψ_k in order to make the numbering $W_0, W_1, \ldots$ acceptable. For $x \in J_k$, let

$$W_x = \begin{cases} \{x+1\} & \text{if } y_k > x+1; \\ \{x+1, x+2\} & \text{if } y_k \leq x+1. \end{cases}$$

It is straightforward to verify that $W_0, W_1, \ldots$ is indeed a numbering, the only important ingredient is that the y_k are approximated from above.

Let $x > 0$ be given and choose the k with $\min(J_k)+1 \leq x \leq \max(J_k)+2$. If $x < y_k$ then $x < \max(J_k)$, $x-1 \in J_k$, $y_k > (x-1)+1$ and $W_{x-1} = \{x\}$, hence $x \in \mathit{NRW}$. If $x > y_k$ then $x \geq \min(J_k)+2$ (as there are three or more numbers $z \in J_k$ with $H(z) \geq k$), $x-2 \in J_k$, $y_k \leq (x-2)+1$ and $W_{x-2} = \{x-1, x\}$, hence $x \in \mathit{NRW}$ again. If $x = y_k$ then $x \in \mathit{NRW}$ only if there is an $\ell \leq k$ with W_{x_ℓ} being nonempty and having the maximum y_k. If this holds, one can compute the value of y_k from k, ℓ by waiting for the first stage s where $y_{k,s} \in W_{x_\ell,s}$; then the value $y_{k,s}$ equals y_k. It follows that $H(y_k) \leq H(k) + H(\ell) + c$ for some constant c independent of k and ℓ. As $H(k)$ and $H(\ell)$ are both logarithmic in k and $H(y_k) \geq k$ for all k, this can happen only for finitely many k. In other words, $y_k \notin \mathit{NRW}$ for almost all k.

This shows that the complement of *NRW* is a finite variant of A. By adjusting the co-2-r.e. approximation to $\overline{A}$ at finitely many places, one receives a co-2-r.e. approximation to *NRW*. □

The next result shows that *NRW* can never by 2-r.e.; as all r.e. and co-r.e. sets are 2-r.e., this result provides a negative answer to Teutsch's question [20].

Theorem 4.2: *For every underlying acceptable numbering φ, the set NRW is not 2-r.e., that is, NRW is not the difference of two r.e. sets.*

Proof: Assume by way of contradiction that $\mathit{NRW} = W_i - W_j$. Note that *NRW* is coinfinite as there are infinitely many indices of the empty set. Using the fixed-point theorem, one can construct sets W_a, W_b such that

$$W_a = \{x : \forall y < x \; [y \in W_i \cup \{0, 1, \ldots, a\}]\};$$
$$W_b = \{x : \exists s \; [x > b \wedge x \in W_{j,s+1} \wedge W_{j,s} \subseteq \{0, 1, \ldots, b\}]\}.$$

If W_j is infinite, then the maximum of W_b is greater than b and is in W_j. But then this maximum is a member of *NRW* but not of $W_i - W_j$. Hence W_j must be finite and W_i coinfinite. So the set W_a is finite as its maximum is the least nonelement of W_i which is greater than a. Again this maximum is a member of *NRW* but not of $W_i - W_j$. This gives a contradiction to $NRW = W_i - W_j$. Hence *NRW* is not a 2-r.e. set. □

In the case that the underlying numbering is a Kolmogorov numbering, a even stronger result can be obtained: *NRW* is not n-r.e. for any natural number n. The proof uses a basic fact stated in the following remark.

Remark 4.3: If the underlying numbering is a Kolmogorov numbering then there is a constant c such that $|\{0, 1, \ldots, cx\} - NRW| > x$ for all x.

To see this, note that $0 \notin NRW$, hence it is enough to look at positive x. Let B_m be the complement of $\{m\}$ for all m. Then there are constants c', d' such that for each B_m equals a set W_n with $n < c'm + d'$. So for every $x > 0$ there are $x + 1$ indices of sets B_m with $m \leq x$ below $(c' + d')x$. Therefore at most $(c' + d')x - x - 1$ of the numbers $0, 1, \ldots, (c' + d')x$ are indices of finite sets. Hence, only $(c' + d')x - x - 1$ of the numbers $0, 1, \ldots, (c' + d')x$ can be in *NRW*. In other words, taking $c = c' + d'$ implies the desired inequality $|\{0, 1, \ldots, cx\} - NRW| > x$ for all x.

Theorem 4.4: *If the underlying numbering is a Kolmogorov numbering then the set NRW is not n-r.e. for any n.*

Proof: Let c be the constant from Remark 4.3. Assume by way of contradiction that there is a number n and an approximation $A_0, A_1, A_2, \ldots$ to *NRW* which makes at most n mind changes for all x. Now one construct a family $B_0, B_1, \ldots$ of sets such that, using the fixed point theorem, one knows the constants c', d' to translate the B_x into some W_y with $y \leq c'x + d'$. Now let $B_{x,0} = \{c'x + d' + 1\}$. At stage s let $y = \max(B_{x,s})$ and check whether the following conditions hold:

- $y \in A_s$;
- $|\{0, 1, \ldots, 3yc\} - A_s| \geq 3y$;
- $\max(B_{u,s}) > 3cy$ for all $u < x$.

If these conditions are satisfied then let $B_{x,s+1} = B_{x,s} \cup \{y + 1\}$ else let $B_{x,s+1} = B_{x,s}$.

If some of the so constructed sets is finite, then choose x to be the least number such that B_x is finite and let $y = \max(B_x)$. As there is a $z \leq c'x + d'$ with $W_z = B_x$ and $y \geq c'x + d' + 1 > z$, it holds that $y = \max(W_z)$ and $y \in NRW$. There is a stage s which is so large that

- $y = \max(B_{x,s})$;

- $A_s(z) = NRW(z)$ for all $z \in \{0, 1, \ldots, 3cy\}$; in particular, $y \in A_s$ and $|\{0, 1, \ldots, 3yc\} - A_s| \geq 3y$.
- for every $u < x$, there is an element larger than $3cy$ of the infinite set B_u enumerated, that is, $\max(B_{u,s}) > 3cy$.

Now it would follow that $B_{x,s+1} = B_{x,s} \cup \{y + 1\}$ by the construction of the set B_x in contradiction to the assumption that $y = \max(B_x)$. Hence B_x has to be infinite.

Given n, let m be so large that every $x \leq 2cn + n + 1$ satisfies the condition $\max(B_{x,0}) < m$. For these x, let s_x be the first stage where $m+1 \in B_{x,s_x+1}$. As $\max(B_{x,s_{x+1}}) > 3mc$ and $|\{0, 1, \ldots, 3mc\} - A_{s_x}| \geq 3m$, there is for any $y \in \{m, m+1, \ldots, 3mc\}$ a stage $t \in \{s_x, s_x+1, \ldots, s_{x+1}-1\}$ with $y = \max(B_{x,t})$ and $y < \max(B_{x,t+1})$, hence $y \in A_t$ for that t. Hence there at least $2m$ numbers $y \in \{m, m + 1, \ldots, 3cm\}$ for which there is a $t \in \{s_x, s_x + 1, \ldots, x_{x+1} - 1\}$ with $A_t(y) = 0 < A_{t+1}(y) = 1$. As this applies to $x = 0, 1, \ldots, 2cn + n$, there are $2m(2cn + n)$ pairs (y, t) with $y \in \{m, m+1, \ldots, 3cm\}$ and $A_t(y) < A_{t+1}(y)$. So there is an number y with $A_t(y) < A_{t+1}(y)$ for at least $2n$ stages t as $|\{m, m+1, \ldots, 3cm\}| = 2cm+1$ and

$$\frac{2m(2cn + n)}{2mc + 1} \geq \frac{2mn(2c + 1)}{(2c + 1)m} = 2n.$$

This contradicts to *NRW* being an n-r.e. set. □

The last result of this paper gives a short alternative proof for Davie's first result that whenever *NRW* is based on a Kolmogorov numbering then $NRW \equiv_T K$. As *NRW* is ω-r.e., obviously $NRW \leq_T K$ and this direction is not included in the proof. After his first result, Davie corresponded with Solovay who indicated to him how to generalize his first result such that it answers Friedman's question for every "usual" underlying acceptable numbering where "usual" means that there exists a polynomial p such that for every recursively enumerable family $A_0, A_1, \ldots$ of r.e. sets and every x there is an $y < p(x)$ with $W_y = A_x$. The proof for Davie's first result given here uses Kummer's Cardinality Theorem [7, 8], but it should be noted that also Owing's preliminary results [17] are sufficient to prove this result as the sets E_n in the proof below are uniformly recursive relative to *NRW*.

Theorem 4.5: *If the underlying numbering is a Kolmogorov numbering then the set NRW is Turing complete; that is,* $NRW \equiv_T K$.

Proof: Note that the constant c from Remark 4.3 satisfies for every x the condition $\{x, x+1, \ldots, cx\} \not\subseteq NRW$. Now let $J_0, J_1, \ldots$ be a partitioning of the natural numbers into intervals satisfying $\min(J_n)c < \max(J_n)$ for all natural numbers n. Furthermore, let $D_0, D_1, D_2, \ldots$ be a canonical indexing of all finite sets of natural numbers with $D_0 = \emptyset$. Define

$$B_m = \{y : y \leq m \cdot (|D_n| + 1) + |K \cap D_n| \text{ for the } n \text{ with } m \in J_n\}$$

and note that the B_m are uniformly r.e. finite sets. Now choose a constant k such that for all n with $|D_n| \geq k$ and all $m \in J_n$, every set B_m has an index u such that $W_u = B_m$ and $u < mk \leq m \cdot (|D_n| + 1) + |K \cap D_n|$. The constant exists since $W_0, W_1, W_2, \ldots$ is a Kolmogorov numbering of r.e. sets. Now one can define the following sets in a way that they are uniformly r.e. relative to *NRW*:

$$E_n = \{u \in \{0, 1, \ldots, k\} : \forall m \in J_n \; [\, |D_n| = k \wedge mk + u \in NRW]\}.$$

By choice, E_n contains $|K \cap D_n|$ whenever $|D_n| = k$. Furthermore, $c(k+1)\min(J_n) < (k+1)\max(J_n)$, thus there is for every n some $u \in \{0, 1, \ldots, k\}$ not contained in E_n. Now applying Kummer's Cardinality Theorem [7, 8] to the cardinality function $\#_k^K$ gives $K \leq_T NRW$. □

Acknowledgments: The author wants to thank Martin Kummer for useful discussions and detailed comments. The author also thanks George Davie, Carl Jockusch, Marcus Schaefer and Jason Teutsch for correspondence; Jason Teutsch permitted him to see an incomplete version of his thesis including the open problems there.

References

1. Gregory J. Chaitin. *A theory of program size formally identical to information theory. Journal of the Association for Computing Machinery* 22:329–340, 1975.
2. Gregory J. Chaitin. Information-theoretic characterizations of recursive infinite strings. *Theoretical Computer Science*, 2:45–48, 1976.
3. Gregory J. Chaitin. Algorithmic information theory, *IBM Journal of Research and Development*, 21:350–359+496, 1977.
4. George Davie. Foundations of Mathematics – Recursion Theory Question. `http://cs.nyu.edu/pipermail/fom/2002-May/005535.html` This posting answers Friedman's Question [6].
5. Rodney G. Downey and Denis R. Hirschfeldt. *Algorithmic Randomness and Complexity.* Manuscript, 2007.
6. Harvey Friedman. Foundations of Mathematics – Recursion Theory Question. `http://cs.nyu.edu/pipermail/fom/2002-February/005289.html`

7. Valentina Harizanov, Martin Kummer and James C. Owings, Jr. Frequency computation and the cardinality theorem. *Journal of Symbolic Logic*, 57:682–687, 1992.
8. Martin Kummer. A proof of Beigel's cardinality conjecture. *The Journal of Symbolic Logic*, 57:677–681, 1992.
9. Martin Kummer. On the complexity of random strings (extended abstract). STACS 1996, *Springer LNCS* 1046:25-36, 1996.
10. Alistair H. Lachlan. On the lattice of recursively enumerable sets. *Transactions of the American Mathematical Society* 130:1–37, 1968.
11. Leonid A. Levin. *Some Theorems on the Algorithmic Approach to Probability Theory and Information Theory*. Dissertation in Mathematics, Moscow, 1971.
12. Ming Li and Paul Vitányi. *An Introduction to Kolmogorov Complexity and Its Applications*. Section 3.5, Second Edition, Springer, 1997.
13. Donald A. Martin. A theorem on hyperhypersimple sets. *The Journal of Symbolic Logic*, 28:273–278, 1963.
14. Andrei A. Muchnik and Semen Ye. Positselsky. Kolmogorov entropy in the context of computability theory. *Theoretical Computer Science* 271:15–35, 2002.
15. Piergiorgio Odifreddi. *Classical Recursion Theory*, *Studies in Logic and the Foundations of Mathematics*, volume 125. North-Holland, Amsterdam, 1989.
16. Piergiorgio Odifreddi. *Classical Recursion Theory II*, *Studies in Logic and the Foundations of Mathematics*, volume 143. Elsevier, Amsterdam, 1999.
17. James C. Owings, Jr. A cardinality version of Beigel's Nonspeedup Theorem. Journal of Symbolic Logic 54:761–767, 1989.
18. Robert W. Robinson. Simplicity of recursively enumerable sets. *The Journal of Symbolic Logic*, 32:167–172, 1967.
19. Robert I. Soare. *Recursively enumerable sets and degrees*. Perspectives in Mathematical Logic. Springer-Verlag, Berlin, 1987.
20. Jason Teutsch. *Noncomputable Spectral Sets*, Question A.5, Draft of forthcoming PhD Thesis, private communication, 2007.
21. C. E. Mike Yates. Recursively enumerable sets and retracing functions. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 8:331–345, 1962.