One-way functions using Algorithmic and Classical Information Theories[*]

Luís Antunes
DCC - FCUP
SQIG at Instituto de Telecomunicações
lfa@dcc.fc.up.pt

Armando Matos
DCC - FCUP
Laboratório de Inteligência artificial e Ciência de Computadores
acm@dcc.fc.up.pt

Alexandre Pinto
ISMAI -
HASLab / INESC TEC
ampinto@docentes.ismai.p

André Souto
DM - IST
SQIG at Instituto de Telecomunicações
asouto@math.ist.utl.pt

Andreia Teixeira
DCC - FCUP
SQIG at Instituto de Telecomunicações
andreiasofia@dcc.fc.up.pt

October 28, 2013

[*] A very preliminary version of this paper appeared in Proceedings of the 6th International Conference on Computability in Europe, Azores, Portugal, July 2010, pages 346-356. The authors are supported by CSI^2 (PTDC/EIA- CCO/099951/2008).

Abstract

We prove several results relating injective one-way functions, time-bounded conditional Kolmogorov complexity, and time-bounded conditional entropy.

First we establish a connection between injective, strong and weak, one-way functions and the expected value of the polynomial time-bounded Kolmogorov complexity, denoted here by $E(K^t(x|f(x)))$. These results are in both directions. More precisely, conditions on $E(K^t(x|f(x)))$ that *imply* that f is a weak one-way function, and properties of $E(K^t(x|f(x)))$ that are *implied* by the fact that f is a strong one-way function. In particular, we prove a separation result: based on the concept of time-bounded Kolmogorov complexity, we find an interval in which every function f is necessarily weak but not strong one-way function.

Then we propose an individual approach to injective one-way functions based on Kolmogorov complexity, defining *Kolmogorov one-way functions* and prove some relationships between the new proposal and the classical definition of one-way functions, showing that a Kolmogorov one-way function is also a deterministic one-way function. A relationship between Kolmogorov one-way functions and the conjecture of polynomial time symmetry of information is also proved.

Finally, we relate $E(K^t(x|f(x)))$ and two forms of time-bounded entropy, the *unpredictable entropy* H^{unp}, in which "one-wayness" of a function can be easily expressed, and the Yao$^+$ entropy, a measure based on compression/decompression schema in which only the decompressor is restricted to be time-bounded.

1 Introduction

Intuitively, a one-way function is a function that is easy to evaluate but hard to invert. The existence of these functions is an open question which implies $\mathbf{P} \neq \mathbf{NP}$. It is well known that the existence of one-way functions is necessary for the existence of pseudo-random generators, digital signatures, identification schemes, and public-key encryption. On the other hand, it is also known [BM84, GMR88, IL89, ILL89, Rom90] that one-way functions are sufficient for the creation of a pseudo-random generator and that trapdoor one-way functions are sufficient for the construction of public-key encryption and signature schemes. Given the importance of one-way functions and the impact of their application, we analyze them at an individual level using Kolmogorov complexity.

Classically, there are several definitions of one-way functions, namely: strong, weak and deterministic. Informally, f is a strong one-way function if all efficient inverting probabilistic algorithms succeed with negligible probability; f is a weak one-way function if all efficient inverting probabilistic algorithms fail with non-negligible probability; in the case of deterministic one-way functions, the function only needs to be resistant to deterministic algorithms that try to invert it. An interesting fact about strong and weak one-way functions is that, although their definitions are not equivalent, weak one-way functions exist if and only if strong one-way functions exist (see [Gol01] for details).

The Kolmogorov complexity, $K(x)$, ([Kol65], [Sol64] and [Cha66]) of an object x is the length of the shortest program producing x in a universal Turing machine. The time-bounded version of Kolmogorov complexity, $K^t(x)$, is the length of the shortest program producing x within time $t(|x|)$.

In this work, we take a fresh look at injective one-way functions using Kolmogorov complexity. Namely, we start by studying the expected value of time-bounded Kolmogorov complexity of an object $x \in \Sigma^n$ given $f(x)$, where f is the description of the function, given by an oracle. Let E denote the expression $E(K^t_f(x|f(x), n))$. We show that if $E > c$ for any positive constant c, then f is a weak one-way function (Theorem 3.1); on the other hand, we show that if f is a strong one-way function, then $E > c \log n$ for every constant c (Theorem 3.2). Based on these results, we introduce a new definition of one-way functions relying on security of individual instances (Definition 3.1) and relate it with the classical notion of weak one-way functions (Corollary 3.4) and with strong one-way functions (Theorem 3.5); we also give a definition of one-way functions based on time-bounded Kolmogorov complexity of the individual instances (Definition 3.2) and a relationship with deterministic one-way functions is studied (Theorem 3.6). For the individual approach, time-bounded Kolmogorov complexity is suitable to study in more detail one-way functions, avoiding few instances that are not secure and giving a more precise measure of security to the other instances. The intuition is that, if x and f are given, then we can compute $f(x)$ in polynomial time; however the converse does not hold, i.e., for the vast majority of x's, given $f(x)$ we cannot compute, in polynomial time, any useful information about x. In fact, we conjecture that the length of a shortest program computing x given $f(x)$, $|x|$, and f should be approximately equal to the length of a shortest program computing x without any auxiliary input.

We define a (t, ε, δ)-secure Kolmogorov one-way function as a function such that the difference between the length of a shortest program computing x, in time t, given $f(x)$ and the length of a shortest program computing x, in time t, without any auxiliary input is smaller than δ with probability greater than ε. We show that for some parameters (t, ε, δ) this new definition is *more restrictive* than weak one-way functions (Corollary 3.4) and that $(t(n), 0, c \log n)$-secure Kolmogorov one-way functions (called Kolmogorov one-way functions) are also more restrictive than deterministic one-way functions (Theorem 3.6).

In [LM93] and [LW95], the authors relate the existence of one-way functions and the conjecture of polynomial time symmetry of information. For the unbounded version of Kolmogorov complexity, symmetry of information was first proved by Levin (as suggested in [ZL70]), but the proof is not valid when polynomial time-bound restrictions are imposed. The conjecture of polynomial time symmetry of information has close connections to several complexity theoretic questions, similar to the connections concerning the existence of one-way functions. In this work, we relate this conjecture with the existence of Kolmogorov one-way functions, by proving that the polynomial time symmetry of information fails if Kolmogorov one-way functions exist (Theorems 4.1 and 4.3).

Various concepts of time-bounded entropy have been used in the literature [Cac97, JA04, GPY09, RW05, Yao82]. They are either based on computational indistinguishability or on efficient compress/decompress schema. In Section 5, we study the relationship between: unpredictable entropy and strong one-way functions (Theorem 5.1), unpredictable entropy and Yao$^+$ (Theorem 5.2) and Yao$^+$ and the time-bounded Kolmogorov complexity (Theorem 5.3).

2 Preliminaries

All strings used are elements of $\Sigma^* = \{0,1\}^*$ and we denote them by x, y, z. The function log denotes the function $\log_2$ and $|.|$ represents the length of a string. The number of elements of a set A is denoted by $\#A$. It is assumed that any time-bound $t(n)$ is constructible and larger than n. We say that $f(n) \in O(g(n))$ iff $\exists k > 0,\ \exists n_0\ \forall n > n_0\, |f(n)| \leq k \cdot |g(n)|$ and that $f(n) \in \omega(g(n))$ iff $\lim\limits_{n\to+\infty} (|f(n)|\,/\,|g(n)|) = +\infty$.

2.1 One-way functions

We present the basic definitions and the results necessary for the rest of this paper.

Definition 2.1 *A function f is* honest *if $|f(x)|$ and $|x|$ are polynomially related, i.e., for some $k > 0$ and for every $x \in \Sigma^\star$,*

$$(|f(x)| \leq |x|^k + k) \wedge (|x| \leq |f(x)|^k + k).$$

In all definitions presented in this paper we assume that f is honest.

Definition 2.2 (Deterministic one-way function) *A function $f : \Sigma^* \to \Sigma^*$ is a* deterministic one-way function *if the following two conditions hold:*

1. *Easy to compute: there is a (deterministic) polynomial time algorithm A such that on every input x, the algorithm A outputs $f(x)$ (i.e., $A(x) = f(x)$).*

2. *Slightly hard to invert: for any deterministic polynomial time algorithm B, for some polynomial $q(\cdot)$, for every sufficiently large n,*

$$\mathrm{pr}_{x\in\Sigma^n}[f(B(f(x),n)) \neq f(x)] > \frac{1}{q(n)}.$$

Definition 2.3 (Weak one-way function) *A function $f : \Sigma^* \to \Sigma^*$ is a* weak one-way function *if the following two conditions hold:*

1. *Easy to compute: there is a (deterministic) polynomial time algorithm A such that on every input x, the algorithm A outputs $f(x)$ (i.e., $A(x) = f(x)$).*

2. *Slightly hard to invert: for any polynomial $t(\cdot)$, there is a polynomial $q(\cdot)$ such that for every probabilistic t-time-bounded algorithm B and for every sufficiently large n,*

$$\mathrm{pr}_{(x,r)\in\Sigma^n\times\Sigma^{t(n)}}[f(B(f(x),r,n))\neq f(x)] > \frac{1}{q(n)}.$$

Definition 2.4 (Strong one-way function) *A function $f:\Sigma^*\to\Sigma^*$ is a* strong one-way function *if the following two conditions hold:*

1. *Easy to compute: there is a (deterministic) polynomial time algorithm A such that on every input x, the algorithm A outputs $f(x)$ (i.e., $A(x) = f(x)$).*

2. *Hard to invert: for any polynomial $t(\cdot)$, for every probabilistic t-time-bounded algorithm B, for every positive polynomial $q(\cdot)$, and for every sufficiently large n,*

$$\mathrm{pr}_{(x,r)\in\Sigma^n\times\Sigma^{t(n)}}[f(B(f(x),r,n)) = f(x)] < \frac{1}{q(n)}.$$

In the previous definitions, r denotes the randomness used by the algorithm B and its length is bounded by its running time.

It is easy to see that any weak one-way function is a deterministic one-way function and that any strong one-way function is a weak one-way function.

2.2 Kolmogorov complexity

Further details on Kolmogorov complexity can be found, for instance, in the comprehensive textbook [LV08]. We will use the prefix-free definition of Kolmogorov complexity. A set of strings A is prefix-free if there are not two strings x and y in A such that x is a proper prefix of y.

Definition 2.5 *Let U be a fixed universal Turing machine with a prefix-free domain. For any strings $x, y \in \Sigma^*$, the* Kolmogorov complexity of x given y with oracle access to f *is:*

$$K_f(x|y) = \min_p\{|p| : U_f(p,y) = x\}.$$

For any time constructible t, the t-time-bounded Kolmogorov complexity of x given y with oracle access to f *is:*

$$K^t_f(x|y) = \min_p\{|p| : U_f(p,y) = x \textit{ in at most } t(|x|) \textit{ steps}\}.$$

The default value for y, the auxiliary input for the program p, is the empty string ε and for oracle f is the null function. In order to avoid overloaded notation, in those cases we typically drop these arguments in the notation. Kolmogorov complexity is machine independent in the sense that we can fix a universal Turing machine U whose program size is at most a constant additive term worse than in any other machine, and the running time is, at most, a logarithmic multiplicative factor slower than in any other machine. One important result in Kolmogorov complexity is the following (see [LV08]).

Theorem 2.1 (Incompressibility Theorem)

1. *For every integer n and for every $x \in \Sigma^n$, $K(x) \leq n + O(\log n)$.*
2. *For every integer n and for every $x \in \Sigma^n$, $K(x|n) \leq n + O(1)$.*
3. *For each r, the set $\{x \in \Sigma^n : K(x) \leq n + K(n) - r\}$ has at most $2^{n-r+O(1)}$ elements.*

In Information Theory, one useful result is the symmetry of information which states that, given two distributions X and Y, $I(X|Y) = I(Y|X)$, where $I(\cdot)$ is the mutual information (see [CT91]). In [ZL70], it is shown that in the resource unbounded case, the symmetry of information concerning the Kolmogorov complexity also holds up to logarithm term:

Theorem 2.2 (Symmetry of Information) *For all strings x and y in Σ^n,*

$$|K(x,y) - K(x) - K(y|x)| \in O(\log n).$$

We will be interested in relating the existence of Kolmogorov one-way functions (Definition 3.2) with polynomial time-bounded symmetry of information.

Hipothesis 2.3 (Polynomial time-bounded symmetry of information) *Let $t(\cdot)$ be a polynomial. For all strings $x, y \in \Sigma^n$,*

$$\left|K^t(x,y) - K^t(x) - K^t(y|x)\right| \in O(\log n).$$

This conjecture is unknown to hold unconditionally, but in [LM93] and in [LW95], it is shown that:

- If $\mathbf{P} = \mathbf{NP}$ then polynomial time symmetry of information holds ([LW95]).
- If deterministic one-way functions exist, then the polynomial time symmetry of information conjecture is false ([LM93, LW95]).

2.3 Time-bounded entropies

There are two forms for each of the following definitions of entropy: parametrized on n, and "global"; in later, the probability distributions are defined over all words of $\Sigma^\star$. It should be clear from the context which version of the definition is being used.

Definition 2.6 (Statistical distance) *The* statistical distance *between two distributions X and Y, denoted by* $\mathrm{dist}(X, Y)$, *is*

$$\max_A |\mathrm{pr}[A(X) = 1] - \mathrm{pr}[A(Y) = 1]|$$

where A is any statistical test (Boolean function). The computational distance *with respect to algorithms running in polynomial time t, denoted by* $\mathrm{cdist}_t(X, Y)$, *limits A to be any algorithm that runs in time t.*

In the following definitions, we assume that ε is non negligible and t is bounded by a polynomial in n. For the significance of these parameters, the reader can consult [LR07] and the references in its bibliography.

Definition 2.7 (Yao entropy, [BSW03]) *For a distribution X, we say that X has* Yao entropy *at least $k(n)$, denoted by $H^{\mathrm{Yao}}_{\varepsilon(n),t(n)}(X) \geq k(n)$, if for every pair of algorithms (c, d), called respectively the "compressor" and the "decompressor" with total running polynomial time $t(n)$,*

$$\mathrm{pr}_{x \in X}\left[d(c(x)) = x\right] \leq 2^{l-k(n)} + \varepsilon(n),$$

where $l = |c(x)|$.

Definition 2.8 (Conditional Yao entropy, [LR07]) *For a distribution (X, Z), we say that X has* Yao entropy at least $k(n)$, conditioned on Z, *denoted by $H^{\mathrm{Yao}}_{\varepsilon(n),t(n)}(X|Z) \geq k(n)$, if for every pair of algorithms (c, d), called respectively the "compressor" and the "decompressor" with total running polynomial time $t(n)$,*

$$\mathrm{pr}_{(x,z) \in (X,Z)}\left[d(c(x, z), z) = x\right] \leq 2^{l-k(n)} + \varepsilon(n),$$

where $l = |c(x, z)|$.

Definition 2.9 (Yao$^+$ entropy, [Pin09]) *It is a variant of Yao entropy, where only d needs to run in polynomial time $t(n)$.*

Definition 2.10 (Conditional Yao$^+$ entropy) *It is a variant of Conditional Yao entropy, where only d needs to run in polynomial time $t(n)$.*

There is another definition of entropy, the "unpredictable entropy" which will be presented in Section 5, see Definition 5.1.

3 One-way functions and Kolmogorov complexity

We present two approaches to define one-way functions using Kolmogorov complexity.

3.1 An expected value approach

We first show how one-way functions are related with the expected value of polynomial time-bounded Kolmogorov complexity over Σ^n. In particular, we show that if the expectation is at least larger than any constant, we have a weak one-way function. On the other hand, we show that if f is a strong one-way function, then the expectation must be larger than logarithmic.

Theorem 3.1 *Let f be an injective and polynomial time computable function. If for every polynomial $t(\cdot)$ and for every constant c, the expected value of $K_f^{t\log t}(x|f(x),r,n)$, over pairs $(x,r) \in \Sigma^n \times \Sigma^{t(n)}$, is larger than c for every sufficiently large n, then f is a weak one-way function.*

Proof. Assume that f is not a weak one-way function. Then, there are a polynomial $t(\cdot)$ and a probabilistic polynomial time algorithm B running in time-bounded by $t(n)$ such that for every polynomial $q(\cdot)$, the algorithm B cannot invert $f(x)$ with negligible probability. In particular, for $q(n) = n^2$, we have $\mathrm{pr}_{(x,r)\in\Sigma^n\times\Sigma^{t(n)}}(B(f(x),r,n) \neq x) \leq 1/n^2$ for infinitely many integers n.

Let $t'(\cdot)$ be a polynomial such that $t'(n) \geq t(n)\log(t(n))$[1]. This time is enough to simulate the algorithm B on the universal Turing machine U (see [LV08], Theorem 7.1.1, page 532). Consider the pairs $(x,r) \in \Sigma^n \times \Sigma^{t(n)}$ that belong to the set $I = \{(x,r) : B(f(x),r,n) = x\}$.

If $(x,r) \in I$ then $K_f^{t'}(x|f(x),r,n) \leq |B| + O(1)$. Thus, for infinitely many $n's$:

$$\begin{aligned}
&E(K_f^{t'}(x|f(x),r,n)) = \sum_{(x,r)\in\Sigma^n\times\Sigma^{t(n)}} \mathrm{pr}(x,r)\cdot K_f^{t'}(x|f(x),r,n) = \\
&= \sum_{(x,r)\in I} \mathrm{pr}(x,r)\cdot K_f^{t'}(x|f(x),r,n) + \sum_{(x,r)\notin I} \mathrm{pr}(x,r)\cdot K_f^{t'}(x|f(x),r,n) \\
&\leq \mathrm{pr}[(x,r)\in I]\cdot(|B| + O(1)) + \sum_{(x,r)\notin I} \mathrm{pr}(x,r)\cdot(n + O(1)) \\
&< |B| + O(1) + \frac{1}{n^2}(n + O(1)).
\end{aligned}$$

Notice that the last line is upper bounded by a constant independent of x. Thus, if $E(K_f^{t'}(x|f(x),r,n)) > c$ is satisfied for all constants c and for every sufficiently large n, then f is a weak one-way function. □

We now present a result that gives some intuition about the expectation of the Kolmogorov complexity of a strong one-way function.

Theorem 3.2 *Let f be an injective and polynomial time computable function. If f is a strong one-way function, then for every constant c and for every polynomial $t(\cdot)$, the expected value of $K_f^t(x|f(x),r,n)$, over pairs $(x,r) \in \Sigma^n \times \Sigma^{t(n)}$, is larger than $c\log n$ for every sufficiently large n.*

[1] We are implicitly using the Linear Speedup Theorem, see [Pap94].

Proof. Assume, for a contradiction, that for some constant c and some polynomial $t(\cdot)$, we have $E(K_f^t(x|f(x),r,n)) \leq c\log n$ infinitely often. Using Markov's inequality we get:

$$\mathrm{pr}_{(x,r)\in\Sigma^n\times\Sigma^{t(n)}}\left[K_f^t(x|f(x),r,n) \leq 2c\log n\right] > 1 - \frac{c\log n}{2c\log n} = \frac{1}{2}. \tag{3.1}$$

We define an algorithm Q that on input $(f(x),r)$ tries to invert $f(x)$, and succeeds for the cases where $K_f^t(x|f(x),r,n) \leq 2c\log n$. This algorithm runs all programs of size up to $2c\log n$ for at most t steps, using the random string r with input $f(x)$. For each such program, Q tests if the output is an inverse of $f(x)$, and if it is, outputs that inverse. If, for the pair (x,r) it happens that $K_f^t(x|f(x),r,n) \leq 2c\log n$, then Q will find a suitable shortest program and output the correct x. Therefore, its success probability is the condition 3.1.

Since there are at most $2^{2c\log n+1} = 2n^{2c}$ programs of length at most $2c\log n$ and each of them runs for a polynomial number of steps, then Q runs in polynomial time. By construction, we know that for infinitely many $n's$,

$$\mathrm{pr}_{(x,r)\in\Sigma^n\times\Sigma^{t(n)}}[Q(f(x),r,n) = x] > \frac{1}{2}.$$

Thus, f is not a strong one-way function. □

Notice that Theorems 3.1 and 3.2 define an interval based on an average value of a polynomial time-bounded Kolmogorov complexity in which every function f is necessarily weak but not strong one-way function.

3.2 An individual approach

The approach to one-way functions proposed in the previous section does not give a satisfactory insight about the security of individual instances of a particular one-way function. In fact, to have an individual instance analysis of security we must have a precise control on the quantity of information that each particular instance may leak. In this section, we give a notion of one-way functions based on Kolmogorov complexity of particular instances.

Definition 3.1 *Let $t(\cdot)$ be some polynomial, $f:\Sigma^n \to \Sigma^m$ an injective and polynomial time computable function and $\delta(\cdot)$ a positive function. We say that an instance x of length n is* (t,δ)-secure *relatively to a random string $r \in \Sigma^{t(n)}$ and to the function f if*

$$K_f^t(x|r,n) - K_f^t(x|f(x),r,n) \leq \delta(n).$$

Let $\varepsilon(\cdot)$ be a function. We say that f is a (t,ε,δ)-secure Kolmogorov one-way function *if for sufficiently large n,*

$$\mathrm{pr}_{(x,r)\in\Sigma^n\times\Sigma^{t(n)}}\left[x \textit{ is not } (t,\delta)\textit{-secure for } r\right] \leq \varepsilon(n).$$

Theorem 3.3 *If f is $(t(n), \varepsilon(n), \delta(n))$-secure Kolmogorov one-way function, then $E(K^t_f(x|f(x), r, n)) \geq (1 - \varepsilon(n)) \cdot (n - \log n - \delta(n)) - 2$.*

Proof. Let $t(\cdot)$ be any polynomial and consider the following sets:

$$\begin{aligned} R^- &= \left\{ (x, r) \in \Sigma^n \times \Sigma^{t(n)} : \begin{array}{l} \left(K^t_f(x|r, n) \leq n - \log n \right) \vee \\ \left(K^t_f(x|r, f(x), n) \leq K^t_f(x|r, n) - \delta(n) \right) \end{array} \right\} \\ R^+ &= \Sigma^n \times \Sigma^{t(n)} \setminus R^- \end{aligned}$$

Using a counting argument, given r, there are less than $2^{n - \log n + 1}$ strings x such that $K^t_f(x|r, n) \leq n - \log n$. Thus, there are at most $2^{n - \log n + 1} \times 2^{t(n)}$ pairs $(x, r) \in \Sigma^n \times \Sigma^{t(n)}$ such that $K^t_f(x|r, n) \leq n - \log n$.

By the assumption of f being a $(t(n), \varepsilon(n), \delta(n))$-secure Kolmogorov one-way function, there are at most $\varepsilon(n) \times 2^{n + t(n)}$ pairs $(x, r) \in \Sigma^n \times \Sigma^{t(n)}$ such that $K^t_f(x|r, f(x), n) \leq K^t_f(x|r, n) - \delta(n)$.

Thus,

$$\begin{aligned} \#R^- &\leq 2^{t(n)} \cdot (2^{n - \log n + 1}) + \varepsilon(n) \cdot 2^{n + t(n)} \\ &= 2^{n + t(n)} \left(2^{-\log n + 1} + \varepsilon(n) \right) \\ &= 2^{n + t(n)} \left(\varepsilon(n) + \frac{2}{n} \right) \end{aligned}$$

The number of pairs in R^+ is at least $\left(1 - \varepsilon(n) - \frac{2}{n}\right) 2^{n + t(n)}$. Thus,

$$\begin{aligned} E(K^t_f(x|f(x), r, n)) &= \sum_{(x,r) \in \Sigma^n \times \Sigma^{t(n)}} \mathrm{pr}(x, r) \cdot K^t_f(x|f(x), r, n) \\ &\geq \sum_{(x,r) \in R^+} \mathrm{pr}(x, r) \cdot K^t_f(x|f(x), r, n) \\ &\geq \left(1 - \varepsilon(n) - \frac{2}{n} \right) \cdot (n - \log n - \delta(n)) \\ &\geq (1 - \varepsilon(n)) \cdot (n - \log n - \delta(n)) - 2 \end{aligned}$$

□

Corollary 3.4 *Let $t(\cdot)$ be a polynomial. If f is $(t(n), \varepsilon(n), \delta(n))$-secure Kolmogorov one-way function with $\varepsilon(n)$ and $\delta(n)$ such that*

$$\lim_{n \to +\infty} ((1 - \varepsilon(n)) \cdot (n - \log n - \delta(n)) - 2) = +\infty,$$

then f is a weak one-way function. In particular, a $(t(n), 1 - \omega(\frac{1}{n}), 0.9n)$-secure Kolmogorov one-way function is a weak one-way function. [2]

Theorem 3.5 *Assume that f is an injective strong one-way function. Then, for every constant c and for every polynomial $t(\cdot)$, f is $(t(n), 1 - 1/n, n - c \log n)$-secure Kolmogorov one-way function.*

[2] When we say that $\varepsilon(n)$ is $1 - \omega(\frac{1}{n})$, we mean that $(1 - \varepsilon(n)) \in \omega(\frac{1}{n})$.

Proof. If f is not $(t(n), 1-1/n, n-c\log n)$-secure Kolmogorov one-way function we have that, for infinitely many $n's$,

$$\mathrm{pr}_{(x,r)\in\Sigma^n\times\Sigma^{t(n)}}(K^t_f(x|r,n) - K^t_f(x|f(x),r,n) \geq n - c\log n) \geq 1 - 1/n$$

which is equivalent to

$$\mathrm{pr}_{(x,r)\in\Sigma^n\times\Sigma^{t(n)}}(K^t_f(x|f(x),r,n) \leq K^t_f(x|r,n) - n + c\log n) \geq 1 - 1/n$$

Consider the following set:

$$I = \{(x,r) \in \Sigma^n \times \Sigma^{t(n)} : K^t_f(x|f(x),r,n) \leq K^t_f(x|r,n) - n + c\log n\}$$

Then,

$$\begin{aligned}
E(K^t_f(x|f(x),r,n)) &= \sum_{(x,r)\in\Sigma^n\times\Sigma^{t(n)}} \mathrm{pr}(x,r)K^t_f(x|f(x),r,n) \\
&= \sum_{(x,r)\in I} \mathrm{pr}(x,r)K^t_f(x|f(x),r,n) + \sum_{(x,r)\notin I} \mathrm{pr}(x,r)K^t_f(x|f(x),r,n) \\
&\leq \sum_{(x,r)\in I} \mathrm{pr}(x,r)K^t_f(x|f(x),r,n) + \sum_{(x,r)\notin I} \mathrm{pr}(x,r)(n+O(1)) \\
&\leq \sum_{(x,r)\in I} \mathrm{pr}(x,r)K^t_f(x|f(x),r,n) + \frac{1}{n}(n+O(1)) \\
&\leq \sum_{(x,r)\in I} \mathrm{pr}(x,r)(c'\log n) + \frac{n+O(1)}{n} \\
&\leq 1\cdot(c'\log n) + 1 + o(n) \leq (c'+2)\log n
\end{aligned}$$

Thus, $E(K^t_f(x|f(x),r,n)) \leq (c'+2)\log n$ and by Theorem 3.2, we conclude that f is not a strong one-way function. □

In order to avoid dealing with probabilities we can think of a different approach based on Definition 3.1.

Definition 3.2 *Let $f : \Sigma^\star \to \Sigma^\star$ be an injective and polynomial time computable function such that $|f(x)| = m(n)$ $\forall x \in \Sigma^n$, where m is some polynomial. We say that f is* Kolmogorov one-way function *if for every polynomial $t(\cdot)$, for every positive integer c, for every sufficiently large n and for every x of length n,*

$$K^t_f(x|n) - K^t_f(x|f(x),n) \leq c\log n.$$

Theorem 3.6 *If f is a Kolmogorov one-way function then f is a deterministic one-way function.*

Proof. We prove this theorem by contraposition. Assume that f is not a deterministic one-way function. Thus, there is a deterministic polynomial time algorithm B such that for every polynomial $q(\cdot)$ and for every n_0, there is an $n \geq n_0$, for which,

$$\#\{x \in \Sigma^n : B(f(x), n) = x\} \geq 2^n - \frac{2^n}{q(n)}.$$

Thus, for an infinity of $n's$, B inverts at least one x such that $|x| = n$, $K_f^t(x|n) > \sqrt{n}$. For these x, we have that $K_f^t(x|n) > \sqrt{n}$ and $K_f^t(x|f(x), n) \leq c'$, where c' is a constant that includes the description of B. Taking those x of sufficiently large n such that for every c, $\sqrt{n} > c \log n + c'$, we have that:

$$\begin{aligned} K_f^t(x|n) - K_f^t(x|f(x), n) \quad & > \sqrt{n} - c' \\ & > c \log n + c' - c' \\ & = c \log n. \end{aligned}$$

□

It is unknown whether the existence of Kolmogorov one-way functions defined as in Definition 3.2 implies the existence of strong or even weak one-way functions.

4 On the Kolmogorov complexity one-way functions and the polynomial time symmetry of information

Longpré and Mocas in [LM93] and Longpré and Watanabe in [LW95] have studied the relationship between classical one-way functions and polynomial time-bounded symmetry of information conjecture.

Similarly, in this section, we explore the connection between the existence of Kolmogorov one-way functions and the polynomial time-bounded symmetry of information. We begin by observing the following:

Theorem 4.1 *If there is a Kolmogorov one-way function with respect to Definition 3.2, then the polynomial time-bounded symmetry of information conjecture does not hold.*

Proof. In Theorem 3.6, we proved that if a Kolmogorov one-way function with respect to Definition 3.2 exists, then a deterministic one-way function also exists. But from [LM93] and [LW95], it is known that if deterministic one-way functions exist then the conjecture of polynomial time-bounded symmetry of information conjecture does not hold. □

Now we introduce the concept of a time function which is used below.

Definition 4.1 *$\tau(n)$ is the smallest function such that:*

$$\exists c, \forall x \in \Sigma^n, K^\tau(x|n) \leq n + c.$$

A function $t(n)$ is a time function *if $t(n) \geq \tau(n)$.*

Notice that $\tau(n)$ exists due to second item of Theorem 2.1 and that $K^t(x)$ is antitone in t.

Lemma 4.2 *Let $t(\cdot)$ be a time function, $k \in \mathbb{N}$ and let $f : \Sigma^n \to \Sigma^m$ be an injective total function such that $|f(x)| = m(n)\ \forall x \in \Sigma^n$, where m is some polynomial. Then, there is a positive constant c such that for all, except a fraction $1/n^k$ of strings x of length n, we have that,*

$$n + c - K^t_f(f(x)|m) \in O(\log n).$$

Proof. As f is injective, a counting argument shows that, for any time function $t(n)$, (more than) $1 - 1/n^k$ of the 2^n strings $f(x)$ have Kolmogorov complexity satisfying $K^t_f(f(x)) \geq n - k\log n$; this holds for any time function $t(n)$. As f is honest, for every x with length n, we have $m = |f(x)| \leq n^i + i$ for some constant integer i. Thus, m cannot "contain" more than $\log m = i\log n + O(1)$ bits of information and we get $K^t_f(f(x)|m) \geq n - O(\log n)$ for at least $(1 - 1/n^k)$ of the strings x with length n. For those strings x, for every time function $t(n)$, and for some positive constant c, we have

$$0 \leq n + c - K^t_f(f(x)|m) \leq (n + c) - (n - O(\log n)) \in O(\log n).$$

□

Notice that this result is also valid for non time-bounded Kolmogorov complexity. As a consequence of this result, we conclude the following.

Theorem 4.3 *For any time function $t(\cdot)$ and for every positive integer k, if the polynomial time-bounded symmetry of information conjecture holds, then $(t(n), \frac{1}{n^k}, O(\log n))$-secure Kolmogorov one-way functions computable in time less than $t(n) - \tau(n)$ do not exist.*

Proof. Let $f : \Sigma^n \to \Sigma^m$ be an injective function computable in polynomial time $t_f(n) \leq t(n) - \tau(n)$ (see Definition 4.1) and let x be a string in Σ^n. For each $r \in \Sigma^{t(n)}$,

1. $K^t_f(f(x)x|r, n, m) \leq K^{t'}_f(x|r, n) + O(1) \leq n + O(1)$, where $t'(n)$ is any time function and $t(n)$ is greater than $t'(n) + t_f(n)$.
2. $n+c-K^t_f(f(x)|r, m) \in O(\log n)$ for all, except a fraction $\frac{1}{n^k}$ of all instances (see Lemma 4.2).

Thus, by polynomial time-bounded symmetry of information, for all, except a fraction $\frac{1}{n^k}$ of all instances, we have

$$\begin{array}{rrcl} & K^t_f(f(x)|r, m) + K^t_f(x|f(x), r, n) & \leq & K^t_f(f(x)x|r, n, m) + O(\log n) \\ \Leftrightarrow & K^t_f(x|f(x), r, n) & \leq & K^t_f(f(x)x|r, n, m) - K^t_f(f(x)|r, m) + O(\log n) \\ \Rightarrow & K^t_f(x|f(x), r, n) & \leq & n + O(1) - K^t_f(f(x)|r, m) + O(\log n), \text{using 1} \\ \Rightarrow & K^t_f(x|f(x), r, n) & \leq & O(\log n) + O(\log n) \in O(\log n), \text{using 2} \end{array}$$

Then, for all, except a fraction $\frac{1}{n^k}$ of all instances and for some constant c,

$$K_f^t(x|r,n) - K_f^t(x|f(x),r,n) \quad \geq K_f^t(x|r,n) - c\log n.$$

Averaging over the strings x with length n and over r, we have that f is not a $(t(n), 1/n^k, O(\log n))$-secure Kolmogorov one-way function, since except for a fraction $\frac{1}{n^k}$ of strings x, we have $K_f^t(x|r,n) \geq \sqrt{n}$ and for those x,

$$K_f^t(x|r,n) - K_f^t(x|f(x),r,n) \geq \sqrt{n} - c\log n > c' \log n$$

for every c' and for sufficiently large n. □

5 One-way functions and time-bounded entropy

In terms of the resources available, the following three concepts, when applied to one-way functions, are somewhat similar.

- The conditional unpredictability entropy $H_{\varepsilon,t}^{\text{unp}}(X|f(X))$, see Definition 5.1 and the reference [LR07].
- The Yao$^+$ entropy $H_{\varepsilon,t}^{\text{Yao}^+}(X|f(X))$, see Definition 2.9 and the reference [Pin09], in which only the decoder algorithm is time-bounded by a polynomial.
- The polynomial time-bounded Kolmogorov ([LV08]) complexity $K^t(x|y)$ in which, again, the expansion time of a minimum program for x is bounded by a polynomial.

We see that, in each case, no time-bounds are imposed to the compressing phase, but the decompressor has to run in polynomial time.

So, it is perhaps not surprising that these concepts are related: in this section, after defining "unpredictability entropy", we present a relationship between one-way functions and a parametrized version of the unpredictability entropy $H^{\text{unp}}(X|f(X))$. Theorem 5.2, which is based on a result of [LR07], compares the *conditional* unpredictability entropy with Yao$^+$ entropy. Theorem 5.3, which is similar to result already published by one of the authors, relates the Yao$^+$ entropy with the average case of a certain time-bounded Kolmogorov complexity. Finally, as a corollary of the results mentioned above we re-obtain a previous result, namely Theorem 3.2.

Definition 5.1 (Unpredictability entropy) *For a distribution (X,Z), we say that X has* unpredictability entropy *at least k conditioned on Z, denoted by $H_{\varepsilon,t}^{\text{unp}}(X|Z) \geq k$, if there is a collection of distributions Y_z (giving rise to a joint distribution (Y,Z)) such that $\text{cdist}_t((X,Z),(Y,Z)) \leq \varepsilon$, and for all probabilistic algorithms running in time t,*

$$E(\text{pr}[A(z,r,n)=x]) \leq 2^{-k(n)}.$$

Let us now suppose that the success probability, as a function of n, decreases faster than the inverse of any positive polynomial (it is called a negligible function), that is, for every positive integer k and for sufficiently large n, $\mathrm{pr}_t(n) \leq 1/n^m$, where $\mathrm{pr}_t(n)$ is the success probability

$$\mathrm{pr}_t(n) = E(\mathrm{pr}[A(z, r, n) = x]).$$

This "negligible" condition can be expressed as: for every positive integer m and for sufficiently large n, $2^{-k(n)} \leq 1/n^m$, that is, $k(n) \geq m \log n$. In other words, $k(n) \in \omega(\log n)$.
Recall now the Definition 2.4 of strong one-way functions. Using Definition 5.1 with the negligible success probability bound, we can easily see that

Theorem 5.1 *A polynomial time computable function $f : \Sigma^n \to \Sigma^n$ is a strong one-way function if the unpredictable entropy $k(n) = H^{\mathrm{unp}}(x|f(x), n) \in \omega(\log n)$.*

The Yao$^+$ entropy is greater than or equal to the unpredictability entropy. This is stated in the following Theorem, whose proof, being analogous to the proof of Lemma 8 in [HLR07], is omitted.

Theorem 5.2 *$H^{\mathrm{unp}}_{\varepsilon,t}(X|Z) \geq k$ implies $H^{\mathrm{Yao}^+}_{\varepsilon,t}(X|Z) \geq k$.*

The following Theorem is similar to a result in [Pin09].

Theorem 5.3 *Let X and Z be probabilistic ensembles, $k(n)$ a function of n, i a positive integer and $t(n)$ a polynomial. If $H^{\mathrm{Yao}^+}_{\varepsilon}(X|Z) \geq k(n) + i$, then $E(K^t(X|Z)) \geq (1 - 2^{-i})k(n)$.*

Proof. Let $c(x, z)$ be a function that, given z, returns the shortest program for x that runs in time $t(|x|)$ on the reference universal Turing machine (UTM); thus, $|c(x, z)| = K^t(x|z)$. Let d be the corresponding decoding function $d(y, z)$, which consists of executing (y, z) in UTM. Define $D_z = \{x \in X : |c(x, z)| < k(n)\}$. Let $c'(x, z)$ be a function similar to $c(x, z)$, but whose output is padded on the right with 0's so that its length is exactly $k(n)$ bits; in order to allow the recovery of $c(x, z)$ from $c'(x, z)$, we are assuming that the UTM is such that minimum programs always end with 1. The respective decoder function $d'(y, z)$ first recovers $c(x, z)$ from $c'(y, z)$ and then executes d on the result $c(x, z)$. We prove the contrapositive of the theorem statement. Suppose that $E(K^t(X|Z)) < (1 - 2^{-i})k(n)$. Then,

$$(1-2^{-i})k(n) > \sum_{x \in X} \mathrm{pr}(x|z) K^t(x|z) \geq \sum_{x \in X \setminus D_z} \mathrm{pr}(x|z)|c(x, z)| \geq k(n) \times \mathrm{pr}[X \notin D_z].$$

So,

$$\mathrm{pr}\,[X \in D_z] > 2^{-i} = 2^{k(n)-(k(n)+i)}.$$

The function d is efficient, because it only has to run a universal Turing machine for a polynomial number of steps. Thus, d' is also efficient, which implies, together with the expression above, that $H^{\mathrm{Yao}^+}_{\varepsilon}(X|Z) < k(n) + i$. □

It has already been observed in [Pin09] that Theorem 5.3 implies the following inequality:

$$H_\varepsilon^{\mathrm{Yao}^+}(X|Z) \leq 2E(K^t(X|Z)). \tag{5.1}$$

It follows from Theorems 5.1, 5.2 and inequality 5.1 that

Theorem 5.4 *If a polynomial time computable function $f : \Sigma^n \to \Sigma^n$ is a strong one-way function, then $E(K^t(x|f(x), r, n)) \in \omega(\log n)$.*

In the proof of Theorem 5.4 a completely different approach is used in order to establish a result equivalent to Theorem 3.2.

6 Conclusions

We used two approaches to relate one-way functions with time-bounded Kolmogorov complexity. In the first, the expected value of $K_f^t(x|f(x), r, n)$ over the strings x with length n has been used as a measure of the difficulty of inverting f; for simplicity, let us denote here that average value by $E(K_f^t(n))$. We believe that Theorem 3.1 is an interesting result, as it shows that a very weak condition, namely $\lim_{n\to\infty} E(K_f^t(n)) = +\infty$, is sufficient to guarantee that a function is a weak one-way function. Together with Theorem 3.2 we obtain a separation result, by defining 2 conditions in terms of $E(K_f^t(n))$ such that, any function satisfying both conditions is necessarily a weak, but not a strong one-way function.

In the second approach, we have tried to individually characterize the "one-way character" of a function. For that purpose, we have defined parametric secure Kolmogorov one-way functions and Kolmogorov one-way functions (Definitions 3.1 and 3.2, respectively). We would like to emphasize here Theorem 3.3, which establishes a general relationship between parametric secure Kolmogorov one-way functions and the expected value of $K_f^t(x|f(x), r, n)$, the measure used in the first approach and, thus, also related to the classical definitions of weak and strong one-way functions. We have also related in Section 4 the conjecture of polynomial time symmetry of information with the non existence of certain parametric secure Kolmogorov one-way functions.

Finally, in Section 5, two forms of conditional time-bounded entropy, namely the Yao$^+$ entropy and the unpredictability entropy were compared with $K_f^t(x|f(x), r, n)$ and, as a corollary, we obtained a very different proof of Theorem 3.2.

7 Acknowledgements

The authors thank to the anonymous reviewers for helpful comments.

The authors also thank EU FEDER and FCT project CSI^2 (PTDC/EIA - CCO/099951 /2008). They are also supported by grants of SQIG -Instituto de Telecomunicações, of HASLab of INESC and FCT grants SFRH/BD/33234/2007 and FRH/BPD/76231/2011.

References

[BM84] M. Blum and S. Micali. How to generate cryptographically strong sequences of pseudo-random bits. *SIAM Journal on Computing*, 13(4):850–864, 1984.

[BSW03] B. Barak, R. Shaltiel, and A. Wigderson. Computational analogues of entropy. In *Proceedings of International Conference on Random Structures and Algorithms*, pages 200–215, 2003.

[Cac97] C. Cachin. Entropy Measures and Unconditional Security in Cryptography. ETH Series in Information Security and Cryptography, Hartung-Gorre Verlag, Konstanz, Germany, 1st edition edition, 1997.

[Cha66] G. Chaitin. On the length of programs for computing finite binary sequences. *Journal of ACM*, 13(4):547–569, 1966.

[CT91] T. Cover and J. Thomas. Elements of information theory. Wiley-Interscience, New York, NY, USA, 1991.

[GMR88] S. Goldwasser, S. Micali, and R. Rivest. A digital signature scheme secure against adaptive chosen-message attacks. *SIAM Journal on Computing*, 17(2):281–308, 1988.

[Gol01] O. Goldreich. Foundations of Cryptography. Cambridge University Press, 2001.

[GPY09] L. Golshani, E. Pasha, and G. Yari. Some properties of Renyi entropy and Renyi entropy rate. *Information Sciences*, 179:2426–2433, June 2009.

[IL89] R. Impagliazzo and M. Luby. One-way functions are essential for complexity based cryptography. In *Proceedings of Symposium on Foundations of Computer Science '89*, 1989.

[ILL89] R. Impagliazzo, L. Levin, and M. Luby. Pseudo-random generation from one-way functions. In *Proceedings of Symposium on Theory of Computing '89*, pages 12–24. ACM, 1989.

[JA04] P. Jizba and T. Arimitsu. The world according to Renyi: Thermodynamics of multifractal systems. *Annals of Physics*, 312:17, 2004.

[Kol65] A. Kolmogorov. Three approaches to the quantitative definition of information. Problems of information transmission, 1(1):1–7, 1965.

[LM93] L. Longpré and S. Mocas. Symmetry of information and one-way functions. *Information Processing Letters*, 46(2):95–100, 1993.

[LR07] C. Lu and L. Reyzin. Conditional Computational Entropy, or Toward Separating Pseudoentropy from Compressibility. *Proceedings of EUROCRYPT'07*, pages 169–186, 2007.

[LV08] M. Li and P. Vitányi. An Introduction to Kolmogorov complexity and its applications. Springer Publishing Company, Incorporated, 2008.

[LW95] L. Longpré and O. Watanabe. On symmetry of information and polynomial time invertibility. *Information and Computation*, 121(1):14–22, 1995.

[Pap94] C. Papadimitriou. Computational Complexity. Addison-Wesley Publucation Company, Reading, MA, 1994.

[Pin09] A. Pinto. Comparing Notions of Computational Entropy. *Theory of Computing Systems*, 45:944–962, 2009.

[Rom90] J. Rompel. One-way functions are necessary and sufficient for secure signatures. In *Proceedings of Symposium on Theory of Computing '90*, pages 387–394. ACM, 1990.

[RW05] R. Renner and S. Wolf. Simple and tight bounds for information reconciliation and privacy amplification. In *Advances in CryptologyASIACRYPT 2005, Lecture Notes in Computer Science*, pages 199–216. Springer-Verlag, 2005.

[Sol64] R. Solomonoff. A formal theory of inductive inference, part I. *Information and Control*, 7(1):1–22, 1964.

[Yao82] A. Yao. Theory and Applications of Trapdoor Functions (Extended Abstract). In *Proceedings of Symposium on Foundations of Computer Science*, pages 80–91, 1982.

[ZL70] A. Zvonkin and L. Levin. The complexity of finite objects and the development of the concepts of information and randomness by means of the theory of algorithms. *Russian Mathematics Surveys*, 256:83–124, 1970.