Kolmogorov complexity and non-determinism

Serge Grigorieff* and Jean-Yves Marion†

Abstract

We are concerned with Kolmogorov complexity of strings produced by non-deterministic algorithms. For this, we consider five classes of non-deterministic description modes : (i) Bounded description modes in which the number of outputs depends on programs, (ii) distributed description modes in which the number of outputs depends on the size of the outputs, (iii) spread description modes in which the number of outputs depends on both programs and the size of the outputs, (iv) description modes for which each string has a unique minimal description, and lastly (v) description modes for which the set of minimal length descriptions is a prefix set.

Keywords : Kolmogorov complexity

Contents

*Université Paris 7, UFR d'Informatique, 2 Place Jussieu, 75251 Paris cedex 05, France - Laboratoire LLAIC, Université Clermont-Ferrand 1 `seg@ufr-info-p7.jussieu.fr`

†LORIA, Université Nancy 2, Calligramme Project, B.P. 239, 54506 Vandœuvre-lès-Nancy Cedex, France, `Jean-Yves.Marion@loria.fr`

1 Introduction : Complexity of Description Modes

Uspensky and Shen in [11] compare various standard definitions of Kolmogorov complexities. For this, they introduce the concept of *description modes*. In essence, a description mode is a binary recursively enumerable (r.e.) relation. So when a description mode turns out to be the graph of a function, it denotes a deterministic computation. Our starting point is to consider description modes in which a program may output more than one string. From this, we attempt to investigate the information content of a string when we deal with non-deterministic algorithms.

In the next subsections, we shall define description modes and their associated complexity measures. Then, we shall give in Subsection 1.5 a full account of each class of description modes that we introduce.

1.1 Description modes and non-determinism

Preliminary Notations 1.

1. Let Δ, Σ both denote the alphabet $\{0, 1\}$. Throughout, Δ^* (resp. Σ^*) denotes words from the alphabet Δ (resp. Σ). The length of a word $\mathtt{p}$ is denoted $|\mathtt{p}|$.

2. We shall use the bijection val : $\Delta^* \to \mathbb{N}$ defined by $\mathrm{val}(\epsilon) = 0$ and $\mathrm{val}(u_0 \cdots u_n) = \sum_{i=0}^{n}(u_i + 1) \cdot 2^i$, where $u_i \in \Delta$ (Caution : this is not the binary development using digits $0, 1$ but the dyadic one which uses digits $1, 2$). We have $|\epsilon| = 0$ and $|u| = \lfloor \log(1 + \mathrm{val}(u)) \rfloor$, for $u \neq \epsilon$, where log denotes the base 2 logarithm.
   Notice that $\mathrm{val}(u) \leq \mathrm{val}(v)$ if u is a prefix or a suffix of v.

3. We shall use the length anti-lexicographic ordering on Δ^*, i.e. length first and then anti-lexicographically, so that $\mathtt{p} < \mathtt{q}$ iff $\mathrm{val}(\mathtt{p}) < \mathrm{val}(\mathtt{q})$.

4. We take as pairing function the function defined by

$$\langle x, y\rangle = x_1x_1x_2x_2\cdots x_nx_n01y \text{ if } x = x_1\cdots x_n.$$

It satisfies the equality $|\langle \mathtt{e}, \mathtt{p}\rangle| = |\mathtt{p}| + 2.|\mathtt{e}| + 2$.

5. Measures of information being considered up to some additive constant, we let $\leq_{\rm ct}$ denote the partial order on functions over natural numbers $\mathbb{N}$ defined by $f \leq_{\rm ct} g$ if there is a constant c such that $f(x) \leq g(x) + c$ for all x.
   Also $f =_{\rm ct} g$ if $f \leq_{\rm ct} g$ and $g \leq_{\rm ct} f$, and $f <_{\rm ct} g$ if $f \leq_{\rm ct} g$ but $g \not\leq_{\rm ct} f$.

6. Throughout, we consider some fixed standard enumeration $(W_{\mathtt{e}})$ of r.e. binary relations included in $\Delta^* \times \Sigma^*$. The set $W_{\mathtt{e}}$ is the domain of the two arguments partial recursive function $\{e\}$.

Refer to [7] for details on plain Kolmogorov complexity.

Description Modes

Following Uspensky and Shen in [10, 11], a *description mode* R is a binary relation on $\Delta^* \times \Sigma^*$ which is r.e. with the following intended meaning. If $(\mathtt{p}, x)$ is in R then we shall say that the program [1] $\mathtt{p}$ generates the string x. Thus, the domain Δ^* is called the set of programs, and the range Σ^* of R is called the set of outputs.

Definition 2. Let R be a description mode and x be a string in Σ^*. The complexity of x is $K_R(x) = \min\{|\mathtt{p}| : (\mathtt{p}, x) \in R\}$.

It must be emphasised that a R-program can generate more than one string. Henceforth, a description mode is naturally constructed as the graph of a non-deterministic computation. To fix thoughts, consider a non-deterministic Turing machine (NDTM) with an output tape. We say that a string x is produced by a NDTM on input $\mathtt{p}$ if x is written on the output tape

[1] We shall use typewriter font for programs

of an accepting computation. Then, define the graph G_M of a NDTM M by $(\mathtt{p}, x) \in G_M$ if there is an accepting computation of $M(\mathtt{p})$ which outputs x. It is clear that G_M is r.e. and so G_M is a description mode. Conversely, from a description mode R, construct M_R as follows. Given $\mathtt{p}$, M_R checks if $(\mathtt{p}, x)$ appears in R. So, the graph M_R is exactly R.

1.2 Optimal Mode and Entropy for a class of modes

Definition 3. According to [10], an additively optimal mode, or in short *optimal mode*, O for a class $\mathcal{C}$ of description modes is a description mode which is in $\mathcal{C}$ and such that $K_O \leq_{\mathrm{ct}} K_R$ for every mode $R \in \mathcal{C}$.

Definition 4. An *entropy* for a class $\mathcal{C}$ is the complexity measure K_O provided by an optimal mode O for $\mathcal{C}$.

Two entropies are clearly equal up to an additive constant. As a consequence, we shall pick up a particular entropy that we shall consider as the entropy of the class $\mathcal{C}$ up to an additive constant.

Classical Kolmogorov (resp. prefix Kolmogorov, cf. Subsection 1.5.6) complexity theory deals with the class of deterministic modes, i.e. graphs of partial recursive functions (resp. with prefix domains). And the theory leans on the existence of optimal modes which lends some credence to intrinsic amount of information inherent to an object.

1.3 Universality

Let $\mathcal{C}$ be a class of description modes.

Definition 5. 1. A description mode $U \in \mathcal{C}$ is a *universal mode* for $\mathcal{C}$ if there is a recursive function $\mathtt{comp} : \Delta^* \times \Delta^* \to \Delta^*$ ($\mathtt{comp}$ stands for compiler) such that, letting $U_{\mathtt{e}} = \{(\mathtt{p}, x) : (\mathtt{comp}(\mathtt{e}, \mathtt{p}), x) \in U\}$, the family $(U_{\mathtt{e}})_{\mathtt{e} \in \Delta^*}$ is an enumeration of the class $\mathcal{C}$.

2. A *strong universal mode* U for $\mathcal{C}$ is an universal mode for $\mathcal{C}$ which satisfies: for each index $\mathtt{e}$, there is a constant c such that for all $\mathtt{p} \in \Delta^*$, we have $|\mathtt{comp}(\mathtt{e}, \mathtt{p})| \leq |\mathtt{p}| + c$.

Proposition 6. *A strong universal mode U for the class $\mathcal{C}$ of description modes is optimal for $\mathcal{C}$.*

Proof. Let S be a description mode in the class $\mathcal{C}$ and let $\mathtt{e}$ be such that $S = U_{\mathtt{e}}$. For all $x \in \Sigma^*$, if $K_S(x)$ is defined then $K_S(x) = |\mathtt{p}|$ with $(\mathtt{p}, x) \in S$. Therefore $(\mathtt{comp}(\mathtt{e}, \mathtt{p}), x) \in U$ and $K_U(x) \leq |\mathtt{comp}(\mathtt{e}, \mathtt{p})| \leq |\mathtt{p}| + c = K_S(x) + c$. Whence, $K_U \leq_{\mathrm{ct}} K_S$. □

Remark 7. A simple way to implement $\mathtt{comp}$ is to use a pairing function, i.e. an injective function $\langle\ ,\ \rangle : \Delta^* \times \Delta^* \to \Delta^*$. Throughout, we take as pairing function the function $\langle x, y\rangle$ defined in Notation 1, Item 4.

Remark 8. Let us make a short digression. Think of a description mode $S \in \mathcal{C}$ as a programming language. An index $\mathtt{e}$ of S may be then considered as an interpreter of S-programs. It turns out that $\lambda p.\, \mathtt{comp}(\mathtt{e}, p)$ is a compiler of S-programs into U-programs, and so $\mathtt{comp}(\mathtt{e}, \mathtt{p})$ is the compiled U-program obtained from the S-program $\mathtt{p}$. (We shall use lambda-notation for functions.) This compiler is based on the interpreter $\mathtt{e}$ of S-programs. In fact, the compiler specialises the interpreter $\mathtt{e}$ to an S-program $\mathtt{p}$ to produce a U-program. That is, it combines both indexes into a suitable program for U. This construction is a very elementary partial evaluation, known as Futamura projection, which is based on the Kleene s^n_m-Theorem. (See the book of Jones [2] for further details.) Now, the function $\lambda e.\, \lambda p.\, \mathtt{comp}(e, p)$, which is obtained by currifying $\mathtt{comp}$, is then a generator of compiler from an interpreter of a description mode in $\mathcal{C}$.

1.4 Effective Universality and Optimality

In practice, $U_{\mathtt{e}}$ is effectively related to $W_{\mathtt{e}}$ and somewhat close to this mode. This leads to the next definition.

Definition 9. 1. U is *max-inclusive universal* for $\mathcal{C}$ if $U \in \mathcal{C}$ and such that for each $\mathtt{e}$, $U_{\mathtt{e}}$ is some maximal (wrt inclusion) submode of $W_{\mathtt{e}}$ lying in $\mathcal{C}$.

2. An *effectively universal mode* U for $\mathcal{C}$ is a universal mode such that for each $\mathtt{e}$ if $W_{\mathtt{e}} \in \mathcal{C}$ then $U_{\mathtt{e}} = W_{\mathtt{e}}$.

3. A strong universal mode U for $\mathcal{C}$ is *effectively strong* in case the constant c in the above definition of strong universality depends recursively on $\mathtt{e}$, i.e. there exists a recursive function $c : \Delta^* \to \mathbb{N}$ such that $\forall \mathtt{p}\, \forall \mathtt{e}\ |\mathtt{comp}(\mathtt{e}, \mathtt{p})| \leq |\mathtt{p}| + c(\mathtt{e})$.

4. U is a *normal universal mode* for $\mathcal{C}$ if it is universal with respect to the function $\mathtt{comp}(\mathtt{e}, \mathtt{p}) = \langle \mathtt{e}, \mathtt{p} \rangle$.

5. O is *effectively optimal* for $\mathcal{C}$ if it is in $\mathcal{C}$ and there exists a recursive function $c : \Delta^* \to \mathbb{N}$ such that $K_O(x) \leq K_R(x) + c(\mathtt{e})$ for every mode $R = W_{\mathtt{e}} \in \mathcal{C}$ and every string $x \in \Sigma^*$.

Notice that the existence of a max-inclusive universal mode implies that every description mode contains a maximal submode lying in $\mathcal{C}$. The following proposition is easy (but useful).

Proposition 10. *1. If U is normal universal then it is effectively strong universal, hence optimal.*

2. If U is max-inclusive universal then it is effectively universal.

3. If U is effectively strong effectively universal (in particular, if U is normal max-inclusive universal) then it is effectively optimal.

Remark 11. 1. It is easy to check that the usual construction of an optimal mode for the class of all deterministic description modes leads to a normal max-inclusive universal mode, hence an effectively optimal mode. The same is true with prefix complexity.
2. All the classes of non deterministic description modes that we shall introduce admit optimal modes, except that of bounded modes. Also, the harmonic classes (cf. Section 6) have optimal modes but no effectively optimal mode.

1.5 Road Map

We shall introduce different classes of description modes in which programs may produce more than one output. Those classes of description modes somehow generalise the classes presented by Uspensky and Shen in [10, 11].

We have to restrict the class of description modes that we shall consider to have a meaningful measure K_R of information. Let us illustrate our intention. Consider the trivial mode $\{0\} \times \Sigma^*$. The complexity of each string is $|0| = 1$. Although this mode is optimal, it says nothing about the information content of a string.

In the next subsections, we shall make an overview of the kind of description modes that we shall investigate. For each kind, we shall give the exact definition, discuss their meanings as computational models, and state the main results. The next Sections will detail the proofs when it is necessary.

1.5.1 Bounded modes

We begin with the study of description modes for which a program outputs a finite number of words.

Definition 12. A *bounded description mode* R is a description mode such that for each program $\mathtt{p}$, the set $\{x : (\mathtt{p}, x) \in R\}$ is finite.

Theorem 13. *The class of bounded description modes has no optimal mode.*

Proof. By contradiction. Suppose that there is an optimal bounded description mode R which provides an entropy K_R. Consider an enumeration of R and let $S = \cup_{k\geq 0} S_k$, where S_k is defined as follows. Suppose that at step t of the enumeration of R, a new pair $(\mathtt{p}, x)$ of R is generated where $|\mathtt{p}| \leq 2k$. Then, we add $(1^k, y)$ to S_k where y is a word which was not produced during the first t steps of the enumeration of R by a program of size $\leq 2k$. Hence, when every pair $(\mathtt{p}, x)$ of R with $|\mathtt{p}| \leq 2k$ has been enumerated (which necessarily happens since R is bounded), we are sure that the last word y added to S_k is not computed by a R program of size $\leq 2k$. Each S_k is a finite subset of $\{1^k\} \times \Sigma^*$ (again because R is a bounded mode). It follows that S is a bounded mode. Now, $\forall k\ \exists y\ K_R(y) \geq K_S(y) + k$. It follows that $K_R \not\leq_{\mathrm{ct}} K_S$ which contradicts the hypothesis. □

Remark 14. The problem of deciding whether or not an index is the code of a bounded description mode is Π^0_3-complete.

1.5.2 β-bounded modes

The above result says that we must restrict more drastically the number of outputs produced by a program, if we seek a notion of entropy. A solution is to bound by a recursive function the number of outputs generated by a program.

Definition 15. Let $\beta : \Delta^* \to \mathbb{N}$ be a recursive function. A *β-bounded mode* R is a description mode R satisfying,

$$\mathrm{card}(\{x \ :\ (\mathtt{p}, x) \in R\}) \leq \beta(\mathtt{p})$$

We shall particularly consider the case where β is *suffix increasing* :

$\beta(\mathtt{p}) \leq \beta(\mathtt{q})$ when $\mathtt{p}$ is a suffix of $\mathtt{q}$ (i.e. $\mathtt{q} = \mathtt{rp}$ for some $\mathtt{r} \in \Delta^*$).

Remark 16. The reason why we deal with suffix increasing functions lies on the standard pairing function, which was defined in Notation 1, Item 4. Though not strictly necessary, it is a convenient hypothesis with which arguments are presented in a clearer way.

Let us give some examples. Consider a NDTM M whose runtime is bounded by $t(n)$ on input of size n. As seen before, M is the graph of some description mode G_M where programs of G_M are considered as inputs of M. Because of the time bound, a program $\mathtt{p}$ shall generate at most $c^{t(|\mathtt{p}|)}$ strings. Put $\beta(\mathtt{p}) = c^{t(|\mathtt{p}|)}$. If t is increasing, then β is suffix increasing. And so, G_M is a β-bounded mode.

In a deterministic mode, each program generates at most one string. So a deterministic mode is a $\lambda x.\, 1$-bounded mode. Conversely, a $\lambda x.\, 1$-bounded mode is deterministic because each program produces at most one string.

We could also see R as a problem. Then, $(\mathtt{p}, x) \in R$ would mean that the instance $\mathtt{p}$ of the problem R has a solution x. Now, saying that $\mathrm{card}(\{x \,:\, (\mathtt{p}, x) \in R\}) \leq \beta(\mathtt{p})$ is equivalent to limit the number of solutions of an instance of a problem. In resource bounded computations, there is an analogous concept which are the counting classes. Of course, this remark is just an analogy, and we don't know what it's worth. But we think it is important to mention it because reasoning by analogy could be fruitful.

In Section 2, we shall establish the existence of optimal β-bounded modes. We shall also relate β-bounded modes to deterministic description modes. Finally, we compare the Kolmogorov complexities associated to the various classes of β-bounded modes and get hierarchy results. As a conclusion to the section, in Subsection 2.6, we consider the case of β-bounded modes with β not necessarily suffix-increasing.

1.5.3 Distributed modes

A description mode R may be also regarded as a class of languages. Indeed, each R-program $\mathtt{p}$ generates the language $R_p = \{x : (\mathtt{p}, x) \in R\}$. Up to now, we have introduced description modes in which each R_p was finite. We now consider the case where each R_p might be infinite. The quantity $K_R(x)$ measures the smallest size of a R-program generating a language which contains x. In other words, $K_R(x)$ is the smallest quantity of information which specifies a property that x satisfies. Similar discussion may be found in [9], in the context of resource bounded computations.

Definition 17. A description mode R is *η-distributed* if for each $\mathtt{p} \in \Delta^*$, $\mathrm{card}\{x : (\mathtt{p}, x) \in R \text{ and } |x| = n\} \leq \eta(n)$ where η is a recursive function.

Notice that the definition above is meaningful when $\eta(n) \leq 2^n$. In fact, each description mode R is a $\lambda n.2^n$-distributed mode because for each $\mathtt{p} \in \Delta^*$, the cardinal of $\{x : (\mathtt{p}, x) \in R \text{ and } |x| = n\}$ is always bounded by 2^n. When η is polynomially increasing, a η-distributed mode R is sparse in the sense that the languages of $(R_{\mathtt{p}})_{\mathtt{p} \in \mathrm{dom}(R)}$ are sparse in the usual sense.

We shall establish in Section 3 that there is an optimal mode for η-distributed modes. We shall examine tradeoffs between information size and the density of distributed modes. Then, we shall establish that the length conditional Kolmogorov entropy $\mathbf{K}_{\mathrm{det}}(x||x|)$ is closely related to the entropies of distributed modes. Finally, we discuss about Loveland uniform entropy and distributed modes.

1.5.4 Spread modes

We now consider a generalisation of both β-bounded modes and η-distributed modes.

Definition 18. Let $\theta : \Delta^* \times \mathbb{N} \to \mathbb{N}$ be suffix increasing with respect to its first argument. A description mode R is a *θ-spread mode* if for each $\mathtt{p} \in \Delta^*$,

$$\mathrm{card}\{x \;:\; (\mathtt{p}, x) \in R \text{ and } |x| = n\} \leq \theta(\mathtt{p}, n)$$

This means, that a program $\mathtt{p}$ of a θ-spread mode may output at most $\theta(\mathtt{p}, n)$ words of length n.

Clearly, a β-bounded mode R is also a θ-spread mode where $\theta(\mathtt{p}, k) = \beta(\mathtt{p})$. Now, take a η-distributed mode S, we see that S is a θ-spread mode where for all $\mathtt{p}$, $\theta(\mathtt{p}, k) = \eta(k)$. In particular, any description mode is a $\lambda\mathtt{p}\lambda k.2^k$-spread mode.

In Section 4, we shall show that there is an entropy for the class of θ-spread modes. Then, we shall establish a hierarchy Theorem for spread modes.

1.5.5 Discriminating modes

A *deterministic mode* D satisfies the following condition : if $(\mathtt{p}, x) \in D$ and $(\mathtt{p}, y) \in D$ then $x = y$. That is, there is a partial recursive function $f : \Delta^* \to \Sigma^*$ such that $(\mathtt{p}, x) \in D$ iff $f(\mathtt{p}) = x$. Deterministic modes are named simple in [11]. The foundation of Kolmogorov Complexity theory is the Invariance Theorem, due to Kolmogorov-Solomonoff.

Theorem 19 (Invariance Theorem). *There is an optimal mode for deterministic modes. The entropy of deterministic modes is denoted* $\mathbf{K}_{\mathrm{det}}$.

Definition 20. Let R be a description mode. The minimal description of a word $x \in \Sigma^*$ in R is

$$E_R(x) = \begin{cases} \min\{\mathtt{p} \;:\; (\mathtt{p}, x) \in R\} & \text{if non empty} \\ \text{undefined} & \text{otherwise} \end{cases}$$

(E_R stands for Exact_R and $<$ is the length anti-lexicographically order on words introduced in Preliminary Notation 1, Item 3.)

Then, $K_R(x) = |E_R(x)|$ is the minimal quantity of information which is necessary to compute x.

A major feature of a deterministic mode is that a program computes a unique string. So, the function E_D is injective if D is a deterministic mode. Notice that geometrically, E_R corresponds to the left contour line of a planar representation of the description mode R.

In general R is not necessarily deterministic. Several strings in Σ^* may have the same minimal description given by E_R. So, those strings are not distinguishable. This leads us to consider the case where E_R is injective.

Definition 21. A *discriminating mode* R is a description mode R for which E_R is injective (though not necessarily total).

Take a discriminating mode R and $\mathtt{p}$ a program. There is at most one x produced by $\mathtt{p}$ such that $E_R(x) = \mathtt{p}$. Hence $E_R(y) < \mathtt{p}$ for every output y of $\mathtt{p}$ which is different from x. In particular, the number of outputs of $\mathtt{p}$ is at most $\mathrm{val}(\mathtt{p}) + 1$.

Proposition 22. *Every discriminating mode is a* $\lambda\mathtt{p}.(1 + \mathrm{val}(\mathtt{p}))$*-bounded mode.*

Remark 23. Given a program $\mathtt{p}$ in the range of E_R and the number n of outputs of $\mathtt{p}$, we can recover the unique string x such that $E_R(x) = \mathtt{p}$. To discriminate x among the outputs $x_0, \cdots, x_n$ of $\mathtt{p}$, it suffices to notice that each x_i, except x has a description smaller than $\mathtt{p}$. Otherwise, the injectivity of E_R would be violated. Henceforth to recover x, we simulate all programs less than $\mathtt{p}$ and wait until $n-1$ strings among $x_0, \cdots, x_n$ are generated. The remaining string is necessarily x.

In Section 5, we shall establish that the class of discriminating modes has an optimal mode which is equivalent, up to an additive constant, to the entropy $\mathbf{K}_{\mathrm{det}}$ of deterministic modes.

1.5.6 Prefix modes

The domain of a mode R is the set $\mathrm{dom}(R) = \{\mathtt{p} \ : \ \exists x \, (\mathtt{p}, x) \in R\}$.

Deterministic prefix modes were introduced by Chaitin [1] and Levin [3] in the context of deterministic computations. We shall consider the non deterministic version.

Definition 24. A deterministic *prefix mode* R is a non deterministic mode for which $\mathrm{dom}(R)$ is prefix-free, i.e. is a set of words such that no word is prefix of another word.

The invariance Theorem for deterministic prefix modes (cf. Chaitin [1], Levin [3]) asserts :

Theorem 25. *There is an optimal mode for the class of deterministic prefix modes. The entropy is denoted* $\mathbf{KP}$*.*

We shall also consider two related conditions

1. *Kraft condition* : $\sum_{\mathtt{p} \in \mathrm{dom}(R)} 2^{|\mathtt{p}|} \leq 1$
2. *harmonic condition* : $\sum_{\mathtt{p} \in \mathrm{dom}(R)} 2^{|\mathtt{p}|}$ is convergent

The issue of Section 6 now is to see what could be a non-deterministic prefix mode with respect to the three prefix-like conditions expressed in Defintions above.

2 β-bounded modes

2.1 Universality and entropy

Theorem 26. *Let β be a suffix increasing recursive function. There is a strong universal mode U for β-bounded modes.*

Proof. We construct an injective recursive function which transforms an index $\mathtt{e}$ of a description mode into an index of a β-bounded mode. Using Kleene s^n_m-Theorem, we construct an injective recursive function $f : \Delta^* \to \Delta^*$ such that the computation of $\{f(\mathtt{e})\}$ on input $(\mathtt{p}, x)$ halts in t steps iff

- $|x| \leq t$ and $|\mathtt{p}| \leq t$ and the computation of $\{\mathtt{e}\}(\mathtt{p}, x)$ halts in t steps,
- $\mathrm{card}\{y \ : \ y < x \textit{ and } \{\mathtt{e}\}(\mathtt{p}, y) \text{ halts in less than } t \text{ steps}\} < \beta(\mathtt{p})$ where $<$ is the length anti-lexicographically order on Σ^* (cf. Notation 1, Item 3).

So, $f(\mathtt{e})$ is the index of a β-bounded mode. Define U as the set of pairs $(\langle f(\mathtt{e}), \mathtt{p}\rangle, x)$ such that $(\mathtt{p}, x) \in W_{f(\mathtt{e})}$. Thus, $U_{\mathtt{e}} = W_{f(\mathtt{e})}$ and $U_{\mathtt{e}}$ is β-bounded. Also, if $\mathtt{e}$ is the index of a β-bounded mode then $W_{f(\mathtt{e})} = W_{\mathtt{e}}$, so that $U_{\mathtt{e}} = W_{f(\mathtt{e})} = W_{\mathtt{e}}$. Thus, the $U_{\mathtt{e}}$'s constitute an enumeration of β-bounded modes.

Now, we show that U is β-bounded. Consider a program $\mathtt{q}$ for U. If $\mathtt{q}$ is not a pair $\langle f(\mathtt{e}), \mathtt{p}\rangle$ then it has no output. If $\mathtt{q}$ is such a pair $\langle f(\mathtt{e}), \mathtt{p}\rangle$ then $\mathtt{e}$ and $\mathtt{p}$ are uniquely determined since f is injective. And the ouptuts of $\mathtt{q}$ for U are exactly those of $\mathtt{p}$ for $W_{f(\mathtt{e})}$, hence their number is at most $\beta(\mathtt{p})$. The choice of the pairing function (cf. Notation 1, Item 4) implies that $\mathtt{p}$ is a suffix of $\langle f(\mathtt{e}), \mathtt{p}\rangle$. Since β is suffix increasing we have $\beta(\mathtt{p}) \leq \beta(\langle f(\mathtt{e}), \mathtt{p}\rangle)$. So U is indeed β-bounded, whence universal for β-bounded modes.

Moreover, since $|\langle f(\mathtt{e}), \mathtt{p}\rangle| \leq |\mathtt{p}| + 2 \cdot |f(\mathtt{e})| + 2$, we see that U is strong universal. □

Corollary 27. *There exists an entropy for β-bounded modes, denoted $\mathbf{K}^{\beta}_{bndd}$.*

Remark 28. It is clear that the above universal mode is normal and max-inclusive, hence effectively optimal.

2.2 An alternative definition

In the previous subsection, we have provided a limitation of description modes from which optimal modes were definable. This limitation was made by uniformly restricting the number of outputs produced by a program. An alternative definition consists in uniformly bounding the number of outputs generated by programs of the same size.

Definition 29. Let $\eta : \mathbb{N} \to \mathbb{R}$ be a function such that the function $\lambda n.2^{\eta(n)}$ is a recursive function from $\mathbb{N}$ to $\mathbb{N}$ and η is strictly increasing. A η-size mode is a description mode R that satisfies, for all n,

$$\mathrm{card}(\{x : (\mathtt{p}, x) \in R \text{ and } |\mathtt{p}| = n\}) \leq 2^{\eta(n)}$$

Say that the volume of computation is the number of all outputs which are produced by all programs of the same size. Following the above definition, the volume of computation of a η-size mode is $\leq 2^{\eta(n)}$. Roughly, it turns out that the volume of computation is the pertinent measure to establish a classification of the β-bounded modes. We first introduce the function $\#\beta$ which measures this volume of computation for β-bounded modes.

Definition 30. Let $\beta : \Delta^* \to \mathbb{N}$ be a suffix increasing function. The function $\#\beta : \mathbb{N} \to \mathbb{R}$ is defined by $\#\beta(n) = \log(\sum_{|\mathtt{p}|=n} \beta(\mathtt{p}))$.

Proposition 31. *If β is a suffix increasing function, then $\#\beta$ is strictly increasing. More precisely, $\#\beta(n+1) \geq \#\beta(n) + 1$ for each n. Hence, $\#\beta(n+c) \geq \#\beta(n) + c$ for every c, and $\#\beta(n-c) \leq \#\beta(n) - c$ if $c \leq n$.*

Proof. For each n, we have

$$\sum_{|\mathtt{p}|=n+1} \beta(\mathtt{p}) = \sum_{i\in\Delta} \sum_{|\mathtt{p}|=n} \beta(i\mathtt{p}) \geq 2 \cdot \sum_{|\mathtt{p}|=n} \beta(\mathtt{p})$$

We obtain $2^{\#\beta(n+1)} \geq 2 \cdot 2^{\#\beta(n)}$. So we have $\#\beta(n+1) \geq \#\beta(n) + 1$. □

The relationship between β-bounded modes and η-size modes is as follows.

Proposition 32. *Let $\beta : \Delta^* \to \mathbb{N}$ be a suffix increasing function. Then*

1. *Each β-bounded mode is a $\#\beta$-size mode*
2. *$\{K_R : R$ is a β-bounded mode$\} = \{K_S : S$ is a $\#\beta$-size mode$\}$*

Proof. (1) Immediate. (2) The left to right inclusion is a consequence of (1). As for the other inclusion, let S be a $\#\beta$-size mode. We define some β-bounded mode R in such a way that the outputs of S-programs of size n are redistributed as outputs of R-programs of size n so that the number of outputs assigned to any R-program $\mathtt{p}$ does not exceed $\beta(\mathtt{p})$. Since the volume of potential computation of R is $\#\beta$, we know that there is enough free space to perform the above construction. Formally, we let $S_n = S \cap (\Delta^n \times \Sigma^*)$ and (J_n) be a recursive sequence of functions such that J_n enumerates S_n without repetition. We construct R as follows : if $J_n(i) = (\mathtt{p}, x)$ and k is such that $\sum_{j<k} \beta(j) \leq i < \sum_{j\leq k} \beta(j)$ (where $\sum_{j<0} \beta(j) = 0$) then put the pair $(\mathtt{q}, x)$ in R where $\mathtt{q}$ is the $k+1$-th program of size n (with respect to the length anti-lexicographic ordering). □

Remark 33. 1. The β-bounded mode R constructed from S in the above proof of item (2) is clearly recursive in S but not many-one reducible to S since it does depend on an enumeration of S.
2. As already said, the key notion is the volume of computation. So, it can

be convenient to consider η-size modes with programs written in unary, i.e. which are a subset of $\{1\}^* \subset \Delta^*$. Indeed, it is not difficult to see that an η-size mode S can be reduced to an η-size mode T with programs in $\{1\}^*$ which has the same complexity function, that is $K_S = K_T$. For this, it suffices to relocate each output of a program of size n of S on the program 1^n of T.

Remark 34. One can also define a function $sum\#\beta : \mathbb{N} \to \mathbb{R}$ as follows :

$$(sum\#\beta)(n) = \log(\textstyle\sum_{|\mathtt{p}|\leq n} \beta(\mathtt{p})) = \log(\sum_{i\leq n} 2^{\#\beta(i)})$$

From Proposition 31 we see that

$$2^{\#\beta(n)} \leq 2^{(sum\#\beta)(n)} \leq (\textstyle\sum_{i\leq n} 2^{i-n}) \times 2^{\#\beta(n)} < 2 \times 2^{\#\beta(n)}$$

Thus, $\#\beta(n) \leq (sum\#\beta)(n) < \#\beta(n) + 1$.

2.3 Information quantity tradeoffs

In order to give a precise relationship between β-bounded modes and deterministic modes in Subsection 2.4 and to establish a hierarchy Theorem for β-bounded modes in Subsection 2.5, we show the following Lemma, which roughly states that the *overall* quantity of information to generate a string is invariant if we switch from α-bounded modes to β-bounded modes. For this, we extend $\#\beta$ to negative numbers by putting $\#\beta(z) = 0$ for all $z < 0$.

Lemma 35. *Let α and β be two suffix increasing functions. There is a constant c such that for all $x \in \Sigma^*$, we have*

$$\#\beta(\mathbf{K}^{\beta}_{bndd}(x) - c) \leq \#\alpha(\mathbf{K}^{\alpha}_{bndd}(x))$$

Proof. Consider an α-bounded mode S. We shall construct a $\#\beta$-size mode T from S such that $\forall x \in \Sigma^*$ $\#\beta(K_T(x) - 1) \leq \#\alpha(K_S(x))$. On the other hand, by Proposition 32, we shall obtain a β-bounded mode R verifying $K_R =_{\mathrm{ct}} K_T$. Next by Corollary 27, we have $\mathbf{K}^{\beta}_{bndd} \leq_{\mathrm{ct}} K_R$. Therefore, we shall conclude that there is c such that for all $x \in \Sigma^*$, we have $\#\beta(\mathbf{K}^{\beta}_{bndd}(x) - c) \leq$

$\#\alpha(K_S(x))$. In particular this inequality holds when S is an optimal α-bounded mode and so the proof will be completed.

Now let us describe the $\#\beta$-size mode T. Define the sequence (I_n) of intervals of integers as follows.

- $I_0 = \{m : 2^{\#\alpha(m)} \leq 2^{\#\beta(0)-1}\}$,
- $I_n = \{m : 2^{\#\beta(n-1)-1} < 2^{\#\alpha(m)} \leq 2^{\#\beta(n)-1}\}$ for $n > 0$

Then define T such that $(1^n, x)$ is in T if there is a pair (q, x) in S such that $|\mathsf{q}| \in I_n$. Clearly, $K_T(x)$ is the unique n such that $K_S(x) \in I_n$.

We verify that T is a $\#\beta$-size mode. By proposition 31, $\#\alpha$ is necessarily strictly increasing. So, we have

$$\sum_{m \in I_n} 2^{\#\alpha(m)} \leq \sum_{0 \leq m \leq m'} 2^{\#\alpha(m)} < 2^{\#\alpha(m')+1}$$

where $m' \in I_n$ is the upper bound of I_n. Next, since by construction $\#\alpha(m') \leq \#\beta(n) - 1$, we obtain that $\sum_{m \in I_n} 2^{\#\alpha(m)} < 2^{\#\beta(n)}$. Therefore, the number of outputs produced by programs of size n of T is bounded by $2^{\#\beta(n)}$. So, T is a $\#\beta$-size mode.

Since $\#\beta$ is strictly increasing, for each $(\mathsf{q}, x) \in S$, there exists n such that $|\mathsf{q}| \in I_n$. When $n > 0$, we have $\#\beta(n-1) \leq \#\alpha(|\mathsf{q}|)$. Taking q such that $|\mathsf{q}| = K_S(x)$ we have $n = K_T(x)$, so that $\#\beta(K_T(x) - 1) \leq \#\alpha(K_S(x))$ for all $x \in \Sigma^*$. $\square$

A very naive approach to Lemma 35 would ask for an improvement to an equality $\#\alpha(\mathbf{K}^{\alpha}_{bndd}(x)) = \#\beta(\mathbf{K}^{\beta}_{bndd}(x))$. But, $\mathbf{K}^{\alpha}_{bndd}$ and $\mathbf{K}^{\beta}_{bndd}$ are defined up to a constant ! Applying a function to $\mathbf{K}^{\alpha}_{bndd}$ really means applying this function to the equivalence class $[\mathbf{K}^{\alpha}_{bndd}]_{=_{\mathrm{ct}}}$. Thus, a less naive approach would ask for an equality

$$\{\#\alpha \circ f : f \in [\mathbf{K}^{\alpha}_{bndd}]_{=_{\mathrm{ct}}}\} = \{\#\beta \circ f : f \in [\mathbf{K}^{\beta}_{bndd}]_{=_{\mathrm{ct}}}\}$$

But $\#\alpha$ and $\#\beta$ may have very different ranges (possibly disjoint) so that again there is no hope for such an equality.
To get an equality from Lemma 35, introduce the partial ordering $\leq$ on functions from $\mathbb{N}$ to $\mathbb{R}$ as $f \leq g$ if for each x, $f(x) \leq g(x)$. Let

$$\mathrm{Final}(X) = \{h : \text{there is } f \in X \text{ such that } h \geq f\}$$

Observe that

$$(*) \quad \forall f \in X\, \exists g \in Y\, g \leq f \text{ is equivalent to } \mathrm{Final}(X) \subseteq \mathrm{Final}(Y)$$

We can reformulate Lemma 35 as follows :

Lemma 36. *The family of functions* $\mathrm{Final}(\{\#\beta \circ f : f =_{\mathrm{ct}} \mathbf{K}^{\beta}_{bndd}\})$ *does not depend on the suffix increasing function* β.

Proof. Lemma 35 can be applied with every $f =_{\mathrm{ct}} \mathbf{K}^{\alpha}_{bndd}$ and insures that there exists $g =_{\mathrm{ct}} \mathbf{K}^{\beta}_{bndd}$ such that $\#\beta \circ g \leq \#\alpha \circ f$. Similarly with α, β exchanged. Then condition (*) insures the equality

$$\mathrm{Final}(\{\#\alpha \circ f : f =_{\mathrm{ct}} \mathbf{K}^{\alpha}_{bndd}\}) = \mathrm{Final}(\{\#\beta \circ f : f =_{\mathrm{ct}} \mathbf{K}^{\beta}_{bndd}\}).$$

□

2.4 Relationship with deterministic modes

A $(\lambda \mathrm{p}.1)$-bounded mode is a mode in which a program outputs at most one string. It follows that the class of $(\lambda \mathrm{p}.1)$-bounded modes is exactly the class of deterministic modes. The volume of computation of deterministic programs of size n is bounded by $\sum_{|p|=n} 1 = 2^n$. So, a deterministic mode is a $(\lambda n.n)$-size mode.
The following result was pointed to us by A. Shen (private communication).

Theorem 37. *There is a constant c such that for all $x \in \Sigma^*$,*

$$\#\beta(\mathbf{K}^{\beta}_{bndd}(x) - c) \leq \mathbf{K}_{\det}(x) \leq \#\beta(\mathbf{K}^{\beta}_{bndd}(x)) + c \leq \#\beta(\mathbf{K}^{\beta}_{bndd}(x) + c)$$

Proof. Set $\alpha = \lambda \mathtt{p}.1$ (whence $\#\alpha = \lambda n.n$) and apply Lemma 35 twice : once as stated and once with α, β exchanged. $\square$

Actually, this relationship can be given a more expressive form. For this, we use the notion of *retract*. Let $f : \mathbb{N} \to \mathbb{R}$ be an injective function. A retract, also named a left inversion, of f is any mapping $g : \mathbb{R} \to \mathbb{N}$ such that $\forall n \in \mathbb{N}\ g(f(n)) = n$.
Notice that a strictly increasing function from $\mathbb{N}$ to $\mathbb{R}$ is injective and in fact we consider throughout only retracts of strictly increasing functions.

Proposition 38. *Let f be a strictly increasing function. Among (non-strictly) increasing retracts of f there is a least one and also a greatest one. The difference between these two extreme monotone retracts is bounded by 1. Hence all monotone retracts of f are equal up to the constant 1. We denote f^{-1} the least monotone retract of f, which is defined as follows :*

for every $r \in \mathbb{R}$, $f^{-1}(r)$ is the greatest $n \in \mathbb{N}$ such that $f(n) \leq r$

Proof. All statements are clear when noticing that the greatest monotone retract g of f is defined by the dual condition

$$g(r) \text{ is the least } n \in \mathbb{N} \text{ such that } f(n) \geq r$$

$\square$

Using Proposition 31, we can now restate Theorem 37 as follows.

Theorem 39. $\mathbf{K}^{\beta}_{bndd} =_{\mathrm{ct}} \#\beta^{-1}(\mathbf{K}_{\mathrm{det}})$

From the classical inequality $\mathbf{K}_{\mathrm{det}}(x) \leq_{\mathrm{ct}} |x|$ we get :

Corollary 40. *There exists a constant c such that for each x*

$$\mathbf{K}^{\beta}_{bndd}(x) \leq \#\beta^{-1}(|x| + c)$$

Constant-bounded modes Finally, consider a mode R in which each R-program may produce at most c outputs. Then, R is a $\lambda n.\, n+\log(c)$-size mode. Another consequence of Lemma 35 is that such a non-deterministic mode R can be translated into a deterministic description mode with an equivalent entropy.

Theorem 41. *For every constant $c > 0$, we have $\mathbf{K}_{\mathrm{det}} =_{\mathrm{ct}} \mathbf{K}^{\lambda p.\, c}_{bndd}$*

The previous result is, in fact, a reformulation of a result of Uspensky and Shen, [11] page 276, that we could loosely rephrase as follows. Suppose that there is a constant c such that for all n, $\mathrm{card}\{x : K_R(x) \leq n\}$ is bounded by 2^{n+c}, where R is a description mode. Then, Uspensky and Shen have proved that $\mathbf{K}_{\mathrm{det}} \leq_{\mathrm{ct}} K_R$.

2.5 An hierarchy theorem for β-bounded modes

In order to compare β-bounded modes, we introduce a partial order on functions.

Definition 42. We denote $\preceq$ the partial order on functions from Δ^* to $\mathbb{N}$ defined as follows.

$$\alpha \preceq \beta \quad \text{if} \quad \exists c \in \mathbb{N}\ \forall n \in \mathbb{N}\ \#\alpha(n) \leq \#\beta(n+c)$$

By $\alpha \simeq \beta$, we mean $\alpha \preceq \beta$ and $\beta \preceq \alpha$, and by $\alpha \prec \beta$, we mean $\alpha \preceq \beta$ but $\beta \not\preceq \alpha$.

For example, take $\alpha(\mathtt{p}) = (|\mathtt{p}| + c)^2$ and $\beta(\mathtt{p}) = |\mathtt{p}|^2$, we can check that $\alpha \preceq \beta$ and so $\alpha \simeq \beta$. Or yet, if $\alpha = O(\beta)$ then we have $\#\alpha \simeq \#\beta$. Geometrically, $\alpha \preceq \beta$ means that $\#\alpha \leq \#\tilde{\beta}$ where $\#\tilde{\beta}$ is obtained by translating $\#\beta$ along the abscissa axis.

Theorem 43. *Let α and β be two suffix increasing functions.*

1. *$\alpha \prec \beta$ iff $\mathbf{K}^{\beta}_{bndd} <_{\mathrm{ct}} \mathbf{K}^{\alpha}_{bndd}$*

2. $\alpha \simeq \beta$ *iff* $\mathbf{K}^{\beta}_{bndd} =_{\mathrm{ct}} \mathbf{K}^{\alpha}_{bndd}$

Proof. Consequence of Lemmas 46 and 48 below. □

Proof of $\alpha \preceq \beta$ **implies** $\mathbf{K}^{\beta}_{bndd} \leq_{\mathrm{ct}} \mathbf{K}^{\alpha}_{bndd}$

Lemma 44. *Let f and g be two strictly increasing functions from $\mathbb{N}$ to $\mathbb{R}$ and c be a positive integer. Assume that for every $n \in \mathbb{N}$, $f(n) \leq g(n+c)$. Then, for every $z \in \mathbb{R}$, $f^{-1}(z) \geq g^{-1}(z) - c$.*

Proof. By definition, we have $f^{-1}(z) = \max(n : f(n) \leq z)$. Since $f(n) \leq g(n+c)$, we get $f^{-1}(z) \geq \max(n : g(n+c) \leq z)$. Now, if $g(c) \leq z$ then $\max(n : g(n+c) \leq z) = \max(t : g(t) \leq z) - c = g^{-1}(z) - c$. If $g(c) > z$ then $g^{-1}(z) < c$ and the wanted inequality is trivial. □

Lemma 45. *Let α and β be two suffix increasing functions such that $\alpha \preceq \beta$. Then, $\#\beta^{-1} \leq_{\mathrm{ct}} \#\alpha^{-1}$*

Proof. Suppose that $\alpha \preceq \beta$. There is c such that for all n, we have $\#\alpha(n) \leq \#\beta(n+c)$. Since $\#\alpha$ and $\#\beta$ are strictly increasing, Lemma 44 yields that the above inequality is equivalent to $\#\alpha^{-1}(z) \geq \#\beta^{-1}(z) - c$ for every $z \in \mathbb{R}$. Consequently, $\#\beta^{-1} \leq_{\mathrm{ct}} \#\alpha^{-1}$ □

Remark that the converse of both Lemmas above holds.

Lemma 46. *Let α and β be two suffix increasing functions such that $\alpha \preceq \beta$. Then, $\mathbf{K}^{\beta}_{bndd} \leq_{\mathrm{ct}} \mathbf{K}^{\alpha}_{bndd}$.*

Proof. By Lemma 35, there is a constant c such that $\#\beta(\mathbf{K}^{\beta}_{bndd}(x) - c) \leq \#\alpha(\mathbf{K}^{\alpha}_{bndd}(x))$. Since $\#\alpha^{-1}$ is increasing, we have

$$\#\alpha^{-1}(\#\beta(\mathbf{K}^{\beta}_{bndd}(x) - c)) \leq \mathbf{K}^{\alpha}_{bndd}(x) \tag{1}$$

Now, since $\alpha \preceq \beta$, Lemma 45 states that there is a constant d such that for every $z \in \mathbb{R}$ $\#\beta^{-1}(z) - d \leq \#\alpha^{-1}(z)$. In particular,

$$\#\beta^{-1}(\#\beta(\mathbf{K}^{\beta}_{bndd}(x)-c))-d \leq \#\alpha^{-1}(\#\beta(\mathbf{K}^{\beta}_{bndd}(x)-c))$$

That is,

$$\mathbf{K}^{\beta}_{bndd}(x)-c-d \leq \#\alpha^{-1}(\#\beta(\mathbf{K}^{\beta}_{bndd}(x)-c))$$

Combining the previous inequality with (1), we get $\mathbf{K}^{\beta}_{bndd} \leq_{\mathrm{ct}} \mathbf{K}^{\alpha}_{bndd}$. □

For example, let $\alpha(x)=\log(|x|)$ and $\beta(x)=|x|$. It is easy to show that $\alpha \preceq \beta$ and so $\mathbf{K}^{\beta}_{bndd} \leq_{\mathrm{ct}} \mathbf{K}^{\alpha}_{bndd}$.

Proof of $\mathbf{K}^{\beta}_{bndd} \leq_{\mathrm{ct}} \mathbf{K}^{\alpha}_{bndd}$ implies $\alpha \preceq \beta$

We first show an incompressibility Lemma for β-bounded modes.

Lemma 47. *There is k such that for each m there is x of size $\lfloor \#\beta(m) \rfloor$ such that*

$$m-k \leq \mathbf{K}^{\beta}_{bndd}(x) \leq m+k$$

Note : In particular, if we consider deterministic mode, we have $\#\beta(m)=m$ and the Lemma says that there is a string x of length m such that $\mathbf{K}_{\mathrm{det}}(x) \geq m-c$, which is a slight weakening of the traditional incompressibility Theorem, as formulated in [7].

Proof. The deterministic incompressibility Theorem asserts that for every n there is a string x of length n such that $\mathbf{K}_{\mathrm{det}}(x) \geq n$. We also know that $\mathbf{K}_{\mathrm{det}}(x) \leq n+c$, for some constant c. Therefore, we have $n \leq \mathbf{K}_{\mathrm{det}}(x) \leq n+c$. Now, fix m and put $n=\lfloor \#\beta(m) \rfloor$ in the inequality above. Hence, there is a string x of size $\lfloor \#\beta(m) \rfloor$ such that

$$\lfloor \#\beta(m) \rfloor \leq \mathbf{K}_{\mathrm{det}}(x) \leq \lfloor \#\beta(m) \rfloor + c$$

where x is of size $\lfloor \#\beta(m) \rfloor$.

Since $\#\beta^{-1}$ is increasing, we also have

$$\#\beta^{-1}(\lfloor \#\beta(m) \rfloor) \leq \#\beta^{-1}(\mathbf{K}_{\mathrm{det}}(x)) \leq \#\beta^{-1}(\lfloor \#\beta(m) \rfloor + c)$$

whence $$\#\beta^{-1}(\#\beta(m)-1) \leq \#\beta^{-1}(\mathbf{K}_{\mathrm{det}}(x)) \leq \#\beta^{-1}(\#\beta(m)+c)$$

Using Proposition 31 and the monotonicity of $\#\beta^{-1}$, we get

$$m - 1 = \#\beta^{-1}(\#\beta(m-1)) \leq \#\beta^{-1}(\#\beta(m) - 1)$$
$$\#\beta^{-1}(\#\beta(m) + c) \leq \#\beta^{-1}(\#\beta(m+c)) = m + c$$

Therefore

$$m - 1 \leq \#\beta^{-1}(\mathbf{K}_{\mathrm{det}}(x)) \leq m + c$$

By Theorem 39, there is a constant d such that for all x,

$$\mathbf{K}^{\beta}_{bndd}(x) - d \leq \#\beta^{-1}(\mathbf{K}_{\mathrm{det}}(x)) \leq \mathbf{K}^{\beta}_{bndd}(x) + d$$

Hence, in the one hand $\mathbf{K}^{\beta}_{bndd}(x) - d \leq m + c$ and on the other $m - 1 \leq \mathbf{K}^{\beta}_{bndd}(x) + d$. Therefore, we see that

$$m - 1 - d \leq \mathbf{K}^{\beta}_{bndd}(x) \leq m + c + d$$

Finally, setting $k = c+d+1$, we conclude that $m-k \leq \mathbf{K}^{\beta}_{bndd}(x) \leq m+k$ □

Lemma 48. *Let α and β be two suffix increasing functions.*

If $\mathbf{K}^{\beta}_{bndd} \leq_{\mathrm{ct}} \mathbf{K}^{\alpha}_{bndd}$ then $\alpha \preceq \beta$.

Proof. Suppose that $\mathbf{K}^{\beta}_{bndd} \leq_{\mathrm{ct}} \mathbf{K}^{\alpha}_{bndd}$. That is, there is a constant c such that $\mathbf{K}^{\beta}_{bndd}(x) \leq \mathbf{K}^{\alpha}_{bndd}(x) + c$ for each $x \in \Sigma^*$. Since $\#\beta$ is increasing, we have $\#\beta(\mathbf{K}^{\beta}_{bndd}(x)) \leq \#\beta(\mathbf{K}^{\alpha}_{bndd}(x) + c)$. Now by Lemma 35, there is a constant d such that $\#\alpha(\mathbf{K}^{\alpha}_{bndd}(x) - d) \leq \#\beta(\mathbf{K}^{\beta}_{bndd}(x))$. By combining both former inequalities, we get

$$\#\alpha(\mathbf{K}^{\alpha}_{bndd}(x) - d) \leq \#\beta(\mathbf{K}^{\alpha}_{bndd}(x) + c) \tag{2}$$

Now, for every m, Lemma 47 claims that there is x such that

$$m - k \leq \mathbf{K}^{\alpha}_{bndd}(x) \leq m + k$$

Since $\#\alpha$ and $\#\beta$ are monotonic, by replacing, in the inequality (2), the two occurences of $\mathbf{K}^{\alpha}_{bndd}(x)$ by $m - k$ and $m + k$ respectively, we obtain $\#\alpha(m - k - d) \leq \#\beta(m + k + c)$. Hence, $\#\alpha(u) \leq \#\beta(u + 2k + c + d)$ for every u, which means $\alpha \preceq \beta$. □

2.6 Relaxing the suffix-increasing hypothesis

As we already observed in Subsection 1.5.2, the hypothesis that β is suffix-increasing is a convenient one but is not really needed to get the diverse results of the preceding subsections. We now look for the exact necessary conditions on β.

First, we consider the existence of a universal mode.

Theorem 49. *Let $\beta : \Delta^* \to \mathbb{N}$ be a recursive function. The class of β-bounded modes contains an universal mode if and only if*

for all $\mathtt{p}$ there are infinitely many $\mathtt{q}$'s such that $\beta(\mathtt{q}) \geq \beta(\mathtt{p})$.

Proof. 1 ($\Rightarrow$). For $\mathtt{r} \in \Sigma^*$ we define $R_{\mathtt{r}}$ as follows :

$$R_{\mathtt{r}} = \{(\mathtt{p}, \langle \mathtt{r}, x \rangle) : \textstyle\sum_{\mathtt{q}<\mathtt{p}} \beta(\mathtt{q}) \leq \mathrm{val}(x) < \sum_{\mathtt{q}\leq\mathtt{p}} \beta(\mathtt{q})\}$$

where val is defined in Item 2 of Notation 1. Then,

1. $\mathrm{card}(\{x : (\mathtt{p}, x) \in R_{\mathtt{r}}\}) = \beta(\mathtt{p})$ for all $\mathtt{p}, \mathtt{r}$
2. $\{x : (\mathtt{p}, x) \in R_{\mathtt{r}}\}$ and $\{x : (\mathtt{q}, x) \in R_{\mathtt{s}}\}$ are disjoint if $(\mathtt{p}, \mathtt{r}) \neq (\mathtt{q}, \mathtt{s})$.

Suppose U is an universal β-bounded mode, relatively to a recursive function $\mathtt{comp}$ and let $\mathtt{e}_{\mathtt{r}}$ be such that $R_{\mathtt{r}} = U_{\mathtt{e}_{\mathtt{r}}}$. Since U is universal, for every $\mathtt{r}, \mathtt{p}$ the program $\mathtt{comp}(\mathtt{e}_{\mathtt{r}}, \mathtt{p})$ has the same outputs for U as the program $\mathtt{p}$ has for $R_{\mathtt{r}}$. Due to Item 1 above and the fact that U is β-bounded, this implies that $\beta(\mathtt{comp}(\mathtt{e}_{\mathtt{r}}, \mathtt{p})) \geq \beta(\mathtt{p})$.

Fix $\mathtt{p}$. Due to Item 2 above, the $\mathtt{comp}(\mathtt{e}_{\mathtt{r}}, \mathtt{p})$'s are distinct when $\mathtt{r}$ varies. Thus, there are infinitely many $\mathtt{q}$'s such that $\beta(\mathtt{q}) \geq \beta(\mathtt{p})$.

2 ($\Leftarrow$). As in the proof of Theorem 26 let f be an injective and recursive function such that $W_{f(\mathtt{e})}$ is a maximal β-bounded mode included in $W_{\mathtt{e}}$. The hypothesis about β allows to define $\mathtt{comp}(\mathtt{e}, \mathtt{p})$ by induction on $\mathrm{val}(\langle \mathtt{e}, \mathtt{p} \rangle)$ as follows : $\mathtt{comp}(\mathtt{e}, \mathtt{p})$ is the first program $\mathtt{q}$ such that $\beta(\mathtt{q}) \geq \beta(\mathtt{p})$ and $\mathtt{q} \neq \mathtt{comp}(\mathtt{e}_1, \mathtt{p}_1)$ for every pair $(\mathtt{e}_1, \mathtt{p}_1)$ such that $\mathrm{val}(\langle \mathtt{e}_1, \mathtt{p}_1 \rangle) < \mathrm{val}(\langle \mathtt{e}, \mathtt{p} \rangle)$.

Define U as the set of pairs $(\mathtt{comp(e,p)}, x)$ such that $(\mathtt{p}, x) \in W_{f(\mathtt{e})}$. The very choice of the function comp shows that U is β-bounded. Also, as in the proof of Theorem 26, we see that the $U_{\mathtt{e}}$'s constitute an enumeration of β-bounded modes. Thus, U is universal for β-bounded modes. □

Remark 50.

1. Theorems 49 together with Theorem 53 below show that there are functions β such that there exists a universal β-mode but no optimal β-mode.

2. A proof similar to that of the above Theorem allows to characterize the β's such that there exists a strong universal β-bounded mode (cf. Definition 5). The (somewhat technical) condition is as follows :

$\forall N\ \exists A\ \forall n\ \exists Q$

$$(Q : N \times \Delta^{\leq n} \to \Delta^{\leq n+A} \text{ is injective } \wedge \forall i\, \forall \mathtt{p}\ \beta(Q(i, \mathtt{p})) \geq \beta(\mathtt{p}))$$

To get the condition for the existence of an optimal mode, we need a lemma, the proof of which is an easy adaptation of that of Proposition 32.

Lemma 51. *Suppose $A \in \mathbb{N}$ and R is such that for all $n \in \mathbb{N}$*

$$\sum_{nA \leq i < (n+1)A} \bigcup_{|\mathtt{p}|=i} \mathrm{card}(\{x : (\mathtt{p}, x) \in R\}) \leq \sum_{nA \leq |\mathtt{p}| < (n+1)A} \beta(\mathtt{p})$$

Then there exists a β-bounded mode S such that $K_S =_{\mathrm{ct}} K_R$.

Yet another lemma relating diverse conditions on β.

Lemma 52. *Let $\beta : \Delta^* \to \mathbb{N}$ be a recursive function. The following conditions are equivalent.*

1. $\forall N\ \exists A\ \forall n\ \sum_{|\mathtt{p}|<n+A} \beta(\mathtt{p}) \geq N \times \sum_{|\mathtt{p}|<n} \beta(\mathtt{p})$

2. $\exists A\ \forall n\ \sum_{n \leq |\mathtt{p}| < n+A} \beta(\mathtt{p}) \geq \sum_{|\mathtt{p}|<n} \beta(\mathtt{p})$

3. $\forall N\ \exists B\ \forall n\ \sum_{(n+1)B \leq |\mathtt{p}| < (n+2)B} \beta(\mathtt{p}) \geq N \times \sum_{nB \leq |\mathtt{p}| < (n+1)B} \beta(\mathtt{p})$

Proof. $1 \Rightarrow 2$. Take $N = 2$ and subtract. $2 \Rightarrow 3$: Apply Condition 2 with $n, n + A, n + 2A, \ldots, n + (N - 1)A$ and set $B = NA$.
$3 \Rightarrow 1$: Applying recursively Condition 3 with 2N yields

$$\sum_{nB \le |\mathtt{p}| < (n+1)B} \beta(\mathtt{p}) \ge (2N)^i \times \sum_{(n-i)B \le |\mathtt{p}| < (n+1-i)B} \beta(\mathtt{p})$$

$$(\sum_{i=1}^{i=n} (2N)^{-i}) \times \sum_{nB \le |\mathtt{p}| < (n+1)B} \beta(\mathtt{p}) \ge \sum_{|\mathtt{p}| < nB} \beta(\mathtt{p})$$

$$\frac{1}{N} \times \sum_{nB \le |\mathtt{p}| < (n+1)B} \beta(\mathtt{p}) \ge \sum_{|\mathtt{p}| < nB} \beta(\mathtt{p})$$

which is condition 1. □

Theorem 53. *Let $\beta : \Delta^* \to \mathbb{N}$ be a recursive function. The class of β-bounded modes contains an optimal mode if and only if β satisfies the equivalent conditions of Lemma 52.*

Proof. 1 ($\Rightarrow$). For $\mathtt{r} \in \Sigma^*$ let $R_\mathtt{r}$ be defined as in the proof of Theorem49 so that

1. $\mathrm{card}(\{x : (\mathtt{p}, x) \in R_\mathtt{r}\}) = \beta(\mathtt{p})$ for all $\mathtt{p}, \mathtt{r}$
2. $\{x : (\mathtt{p}, x) \in R_\mathtt{r}\}$ and $\{x : (\mathtt{q}, x) \in R_\mathtt{s}\}$ are disjoint if $(\mathtt{p}, \mathtt{r}) \neq (\mathtt{q}, \mathtt{s})$.

Suppose O is an optimal β-bounded mode. Consider N many distinct strings $\mathtt{r}_1, \ldots, \mathtt{r}_N$. Since O is optimal there is a constant A such that

(*) $K_O \le K_{R_{\mathtt{r}_i}} + A$ for $i = 1, \ldots, N$.

Fix n. Condition (*) means that for every $\mathtt{p}, i$ such that $|\mathtt{p}| \le n$ and $1 \le i \le N$ each output of program $\mathtt{p}$ for $R_{\mathtt{r}_i}$ is an output for O of some program $\mathtt{q}$ such that $|\mathtt{q}| \le n + A$. Due to Item 1 and 2 above, this implies that

$$N \times \sum_{|\mathtt{p}| < n} \beta(\mathtt{p}) \le \mathrm{card}(\bigcup_{|\mathtt{q}| < n+A} \{x : (\mathtt{q}, x) \in O\})$$

Since O is β-bounded the right member is $\leq \sum_{|\mathtt{q}|<n+A} \beta(\mathtt{q})$, which gives Condition 1 of Lemma 52.

2 ($\Leftarrow$). As in the proof of Theorem 26 let f be an injective and recursive function such that $W_{f(\mathtt{e})}$ is a maximal β-bounded mode included in $W_{\mathtt{e}}$. Suppose A is as in condition 2 of Lemma 52 above and set

$$\mathtt{comp}(\mathtt{e},\mathtt{p}) = 0^{A\times \mathrm{val}(\mathtt{e})}\ 1^A\ \mathtt{p}$$
$$U = \{(\mathtt{comp}(\mathtt{e},\mathtt{p}), x) : (\mathtt{p}, x) \in W_{f(\mathtt{e})}\}.$$

Clearly, $\mathtt{comp}$ is injective. As in the proof of Theorem 26, we see that the $U_{\mathtt{e}}$'s constitute an enumeration of β-bounded modes.
Also, by very construction we have $K_U(x) \leq K_{W_{f(\mathtt{e})}}(x) + A(\mathrm{val}(\mathtt{e})+1)$ for all x. Thus,

(*) $K_U(x) \leq_{\mathrm{ct}} K_R$ for every β-bounded mode R.

We now show that U satifies the condition of Lemma 51. In fact,
$\sum_{nA\leq|\mathtt{q}|<(n+1)A} \mathrm{card}(\{x : (\mathtt{q}, x) \in U\})$
$= \sum_{nA\leq|\mathtt{p}|+A(\mathrm{val}(\mathtt{e})+1)<(n+1)A} \mathrm{card}(\{x : (\mathtt{comp}(\mathtt{e},\mathtt{p}), x) \in U\})$
$\leq \sum_{nA\leq|\mathtt{p}|+A(\mathrm{val}(\mathtt{e})+1)<(n+1)A} \beta(\mathtt{p}) = \sum_{i=0}^{i=n-1} \sum_{(n-i-1)A\leq|\mathtt{p}|<(n-i)A} \beta(\mathtt{p})$
$= \sum_{|\mathtt{p}|<nA} \beta(\mathtt{p}) \leq \sum_{nA\leq|\mathtt{q}|<(n+1)A} \beta(\mathtt{p})$ by condition 1 of of Lemma 52. Thus, we can apply Lemma 51 and get a β-bounded mode O such that $K_O =_{\mathrm{ct}} K_U$. Using property (*) above, we see that O is optimal for β-bounded modes. □

Remark 54. 1. Entropies associated to suffix-increasing functions β are exactly those associated to functions β satisfying condition 3 of Lemma 52 with $N = 2$ and the constant $B = 1$. In fact, suppose this condition is true with $B = 1$ and let γ be defined as follows :

$$\text{if } |\mathtt{p}| = n \text{ then } \gamma(\mathtt{p}) = \frac{\sum_{|\mathtt{q}|=n} \beta(\mathtt{p})}{2^n}$$

(i.e. we equidistribute the outputs on programs having the same length). Condition 3 of Lemma 52 with $N = 2$, $B = 1$ insures that γ is suffix-increasing. Also, the proof of Proposition 32 easily adapts. In particular, $\mathbf{K}^{\beta}_{bndd} = \mathbf{K}^{\gamma}_{bndd}$.

2. Up to now, we always suppposed β to be recursive. One can weaken this hypothesis to β is recursively enumerable from below, i.e. there exists a recursive (non strictly) increasing sequence of functions $(\beta_n)_{n\in\mathbb{N}}$ such that $\beta(\mathtt{p}) = \sup\{\beta_n(\mathtt{p}) : n \in \mathbb{N}\}$. All results go through with no problem. The interest of such an extension lies in the fact that the function which associates to any program $\mathtt{p}$ the number of its outputs for a mode R is not recursive in the general case but is always recursively enumerable from below.

3 Distributed modes

3.1 Universal Distributed modes

Theorem 55. *There is a strong universal mode for η-distributed modes (cf. Definition 17).*

Proof. Suppose that $\mathtt{e}$ is the index of a description mode. We define an injective recursive function $f : \Delta^* \to \Delta^*$ which transforms $\mathtt{e}$ into an index of an η-distributed mode. For this, the program $f(\mathtt{e})$ checks, by dovetailing, if $(\mathtt{p}, x) \in W_{\mathtt{e}}$ for every x such that $|x| = n$. If, during this process, the computation of $\{\mathtt{e}\}(\mathtt{p}, x)$ halts and no more than $\eta(n)$ other computations were terminated, then it outputs 1, i.e. $(\mathtt{p}, x) \in W_{f(\mathtt{e})}$, otherwise it diverges.

Define U as the set of pairs $(\langle f(\mathtt{e}), \mathtt{p}\rangle, x)$ such that $(\mathtt{p}, x) \in W_{f(\mathtt{e})}$. Thus, $U_{\mathtt{e}} = W_{f(\mathtt{e})}$ and $U_{\mathtt{e}}$ is η-distributed. Also, if $\mathtt{e}$ is the index of an η-distributed mode then $W_{f(\mathtt{e})} = W_{\mathtt{e}}$, so that $U_{\mathtt{e}} = W_{f(\mathtt{e})} = W_{\mathtt{e}}$. Thus, the $U_{\mathtt{e}}$'s constitute an enumeration of η-distributed modes.

Now, we show that U is η-distributed. Consider a program $\mathtt{q}$ for U. If $\mathtt{q}$ is not a pair $\langle f(\mathtt{e}), \mathtt{p}\rangle$ then it has no output. If $\mathtt{q}$ is such a pair $\langle f(\mathtt{e}), \mathtt{p}\rangle$ then $\mathtt{e}$ and $\mathtt{p}$ are uniquely determined since f is injective. And the ouptuts of $\mathtt{q}$ for U are exactly those of $\mathtt{p}$ for $W_{f(\mathtt{e})}$. In particular, there are no more than $\eta(n)$ outputs with length n since $W_{f(\mathtt{e})}$ is η-distributed. So U is indeed η-distributed, whence universal. $\square$

As usual, the strong universal mode for η-distributed modes constructed above (is normal max-inclusive universal and) provides an (effectively) optimal mode and so an entropy.

Corollary 56. *There is an entropy $\mathbf{K}^{\eta}_{\text{distri}}$ for the class of η-distributed modes.*

3.2 An hierarchy theorem for distributed modes

Finally, we prove the following relationship between distributed modes.

Theorem 57. *For all $x \in \Sigma^*$ such that $\theta(|x|) > 0$ and $\eta(|x|) > 0$*

$$\mathbf{K}^{\theta}_{\text{distri}}(x) + \log(\theta(|x|)) =_{\text{ct}} \mathbf{K}^{\theta}_{\text{distri}}(x) + \log(\eta(|x|))$$

Proof. Consider a η-distributed mode R. We construct a θ-distributed mode S as follows. The S-program $\mathtt{p}$ produces all words of size n which are the outputs of R-programs whose associated numerical values are in

$$[\text{val}(\mathtt{p}) \cdot \tfrac{\theta(n)}{\eta(n)}, (\text{val}(\mathtt{p}) + 1) \cdot \tfrac{\theta(n)}{\eta(n)}[$$

So, if $(\mathtt{q}, x) \in R$, there is $(\mathtt{p}, x) \in S$ such that $\mathtt{p} = \lfloor \text{val}(\mathtt{q}) \cdot \frac{\eta(|x|)}{\theta(|x|)} \rfloor$.
This implies that $K_S(x) \leq K_R(x) - (\log(\theta(|x|)) - \log(\eta(|x|)))$. By considering an optimal R we conclude that . $\mathbf{K}^{\theta}_{\text{distri}}(x) \leq_{\text{ct}} \mathbf{K}^{\theta}_{\text{distri}}(x) - (\log(\theta(|x|)) - \log(\eta(|x|)))$. By symmetry, we get the wanted equality. $\square$

Remark 58. If $\eta(|x|) = 0$ then an η-distributed mode has no length n output, so that $\mathbf{K}^{\theta}_{\text{distri}}(x) = +\infty$ for every $x \in \Sigma^n$, while $\log(\eta(|x|)) = -\infty$.

3.3 Loveland uniform entropy

It is worth noticing that an example of distributed modes is given by the *uniform entropy* proposed by Loveland in [8]. Again, we follow the presentation given by [11] to present the uniform entropy. (Although, we are not following their terminology where the uniform entropy is called decision entropy.) A description mode R is uniform if

- If $(\mathtt{p}, x) \in R$ then for each $y <_{\text{prefix}} x$, $(\mathtt{p}, y) \in R$.
- If $(\mathtt{p}, x) \in R$ and $(\mathtt{p}, y) \in R$ then either $y \leq_{\text{prefix}} x$ or $x <_{\text{prefix}} y$.

where $<_{\text{prefix}}$ is the prefix order on Σ^*. Clearly, uniform modes form a subclass of $\lambda n.1$-distributed modes. Loveland has shown that there is an optimal mode for uniform description modes.

We can combine both properties, i.e. consider description modes which are both distributed and uniform.

Definition 59. A description mode R is η-distributed uniform if R is η-distributed and if $(\mathtt{p}, x) \in R$ then for each $y <_{\text{prefix}} x$, $(\mathtt{p}, y) \in R$.

Theorem 60. *There is a universal mode for η-distributed uniform modes.*

Proof. Let U be the universal and optimal mode which is described in the proof of Theorem 55. Define the mode U' as the r.e. mode which satisfies the following condition :

$$(\mathtt{q}, x) \in U' \quad \text{iff} \quad \forall y \leq_{\text{prefix}} x, (\mathtt{q}, y) \in U$$

U' is η-distributed as is U. The uniformity condition is implied by the very construction. It follows that U' is an optimal mode for the η-distributed uniform modes. □

3.4 Distributed modes and length conditional complexity

It is shown in [8, 11] that Loveland uniform entropy (see Definition 59) is strictly less than the deterministic entropy $\mathbf{K}_{\rm det}$, with respect to $<_{\rm ct}$. As a consequence, each entropy $\mathbf{K}^{\eta}_{\rm distri}$ is strictly less than the deterministic entropy $\mathbf{K}_{\rm det}$, i.e. $\mathbf{K}^{\eta}_{\rm distri}<_{\rm ct}\mathbf{K}_{\rm det}$.

Now, we compare the length conditional Kolmogorov entropy $\mathbf{K}_{\rm det}(x||x|)$ with entropies of distributed modes. In our setting, the $\mathbf{K}_{\rm det}(x||x|)$ might be defined as follows. Consider a r.e. ternary relation T on $\Delta^*\times\mathbb{N}\times\Sigma^*$. Say that T is a length conditional mode if there is a partial recursive function h such that $(\mathtt{p},n,x)\in T$ iff $h(\mathtt{p},n)=x$ and $|x|=n$. We readily adapt the notion of strong universal mode (and its existence proof) to length conditional modes. Hence, $\mathbf{K}_{\rm det}(x||x|)=\min\{|\mathtt{p}| \ : \ (\mathtt{p},n,x)\in T\}$ where T is a strong universal mode for length conditional modes.

Theorem 61. *1.* $\mathbf{K}^{\eta}_{\rm distri}(x)+\log(\eta(|x|))=_{\rm ct}\sup(\mathbf{K}_{\rm det}(x||x|),\log(\eta(|x|)))$ *for all* $x\in\Sigma^*$ *such that* $\eta(|x|)\neq 0$.

2. $\mathbf{K}^{\eta}_{\rm distri}(x)=_{\rm ct}\sup(0,\mathbf{K}_{\rm det}(x||x|)-\log(\eta(|x|)))$ *for all* $x\in\Sigma^*$.

Proof. Item 1 is a reformulation of Item 2.
First, we establish that $\mathbf{K}_{\rm det}(x||x|)\leq_{\rm ct}\mathbf{K}^{\eta}_{\rm distri}(x)+\log(\eta(|x|))$. Consider an enumeration of a η-distributed mode R. Define the ternary partial recursive function f as follows. $f(\mathtt{p},i,n)=x$ iff $|x|=n$ and i is the rank (relative to the enumeration) of pair $(\mathtt{p},x)$ among R-pairs $(\mathtt{p},y)$ where $|y|=n$. Since R is a η-distributed mode, we have $i\leq\eta(n)$. Define $\mathrm{enc}(i)$ as the binary expansion of i of length $\lceil\log(\eta(n))\rceil$. We can encode the pair $(\mathtt{p},i)$ by merely concatenating $\mathrm{enc}(i)$ and $\mathtt{p}$. Indeed from n and a program which computes the function η, we can recover i and $\mathtt{p}$ from the word $\mathrm{enc}(i)\mathtt{p}$.
Let T be the length conditional mode defined by $g(\mathrm{enc}(i)\mathtt{p},n)=f(\mathtt{p},i,n)$. Since $|\mathrm{enc}(i)\mathtt{p}|\leq|\mathrm{enc}(i)|+|\mathtt{p}|\leq|\mathtt{p}|+\log(\eta(|x|))$ we get

$$K_T(x||x|) \leq_{\mathrm{ct}} K_R(x) + \log(\eta(|x|))$$

Taking R optimal, this yields

$$K_T(x||x|) \leq_{\mathrm{ct}} \mathbf{K}^{\eta}_{\mathrm{distri}}(x) + \log(\eta(|x|))$$

Whence

$$\mathbf{K}_{\mathrm{det}}(x||x|) \leq_{\mathrm{ct}} \mathbf{K}^{\eta}_{\mathrm{distri}}(x) + \log(\eta(|x|))$$

Now, we establish the reverse inequality. For this, consider a length conditional mode T. We define a η-distributed mode R as follows.

The pair $(\mathtt{q}, x)$ is in R iff there is a program $\mathtt{p}$ such that

$$\mathrm{val}(\mathtt{p}) \in [\mathrm{val}(\mathtt{q}) \cdot \eta(|x|), (\mathrm{val}(\mathtt{q}) + 1) \cdot \eta(|x|)[\text{ and } (\mathtt{p}, |x|, x) \in T$$

If $\mathrm{val}(\mathtt{q}) \neq 0$, we have $|\mathtt{q}| \leq_{\mathrm{ct}} |\mathtt{p}| - \log(\eta(|x|))$.

It follows that $|\mathtt{q}| \leq_{\mathrm{ct}} \sup(0, |\mathtt{p}| - \log(\eta(|x|)))$.

We conclude that $K_R \leq_{\mathrm{ct}} \sup(0, K_T(x||x|) - \log(\eta(x)))$. Taking T optimal leads to the wanted inequality. □

Remark 62. 1. Observe that $\mathbf{K}_{\mathrm{det}}(x||x|) - \log(\eta(|x|)) < 0$ does happen. For instance, with $\eta(n) = 2^n$.

2. Cf. also Remark 58.

4 Spread modes

In this section, we present results on spread modes (cf. Definition18). Actually, most of the proofs can be readily adapted from proofs in Section 2, so we shall just sketch them.

Theorem 63. *Let $\theta : \Delta^* \times \mathbb{N} \to \mathbb{N}$ be a recursive function which is suffix-increasing with respect to its first argument. There is a strong universal θ-spread mode. We denote $\mathbf{K}^{\theta}_{\mathrm{spread}}$ the associated entropy.*

Proof. For each $\mathtt{p} \in \Delta^*$, replace $\eta(n)$ in the proof of Theorem 55 by $\theta(\mathtt{p}, n)$. □

In order to compare θ-spread modes with deterministic modes and further to state a hierarchy Theorem, we establish a result which is analogous to Lemma 35. Let $\#\theta(n,k) = \log(\sum_{|\mathtt{p}|=n} \theta(\mathtt{p},k))$.

Lemma 64. *There is a constant c such that for all $x \in \Sigma^*$, we have*

$$\#\theta_2(\mathbf{K}^{\theta_2}_{\text{spread}}(x) - c) \leq \#\theta_1(\mathbf{K}^{\theta_1}_{\text{spread}}(x))$$

Proof. Suppose that S is a θ_1-spread mode. Fix k and define S_k as the description mode which contains all pairs $(\mathtt{p}, x)$ of S where $|x| = k$. Actually, S_k is a α_k-bounded mode where $\alpha_k(\mathtt{p}) = \theta_1(\mathtt{p}, k)$. Now, define $\beta_k : \Delta^* \to \mathbb{N}$ as the suffix increasing function satisfying $\beta_k(\mathtt{p}) = \theta_2(\mathtt{p}, k)$. We can apply the construction of the proof of Lemma 35 to obtain a β_k-bounded mode T_k from the α_k-bounded mode S_k such that $\#\beta_k(K_{T_k}(x) - 1) \leq \#\alpha_k(K_{S_k}(x))$, for every $x \in \Sigma^*$.

We construct a θ_2-spread mode T from the sequence of β_k-bounded modes T_k's. A string x of length k is produced by a program $\mathtt{p}$ of T (i.e. $(\mathtt{p}, x) \in T$) iff $(\mathtt{p}, x) \in T_k$. It follows that $\#\beta_k(K_T(x) - 1) \leq \#\alpha_k(K_S(x))$. T is a θ_2-spread mode because $\beta_k(\mathtt{p}) = \theta_2(\mathtt{p}, k)$ for each $\mathtt{p}$ and so we have $\#\theta_2(K_T(x) - 1) \leq \#\theta_1(K_S(x))$. We conclude as in the proof of Lemma 35. $\square$

Since deterministic modes are surely $\lambda\mathtt{p}\lambda k.1$-spread modes, the above Lemma yields

Corollary 65. *There is c such that for all x,*

$$\#\theta(\mathbf{K}^{\theta}_{\text{spread}}(x) - c, |x|) \leq \mathbf{K}_{\text{det}}(x) \leq \#\theta(\mathbf{K}^{\theta}_{\text{spread}}(x), |x|) + c \leq \#\theta(\mathbf{K}^{\theta}_{\text{spread}}(x) + c, |x|)$$

Let $f : \mathbb{N} \times \mathbb{N} \to \mathbb{R}$ be strictly increasing function with respect to its first variable. We define $f^{-1} : \mathbb{R} \times \mathbb{N} \to \mathbb{N}$ as the function such that for every $r \in \mathbb{R}$,

$$f^{-1}(r,k) \text{ is the greatest } n \in \mathbb{N} \text{ such that } f(n,k) \geq r$$

Notice that $f^{-1}(f(n,k),k) = n$.

Theorem 66.

$$\mathbf{K}^{\theta}_{\text{spread}}(x) =_{\text{ct}} \#\theta^{-1}(\mathbf{K}_{\det}(x), |x|)$$

We extend the ordering $\preceq$ to θ functions as follows.

$$\theta_1 \preceq \theta_2 \quad \text{if} \quad \exists c \in \mathbb{N}\, \forall n \in \mathbb{N}\ \ \forall k \in \mathbb{N}\ \ \#\theta_1(n,k) \leq \#\theta_2(n+c,k)$$

Theorem 67. *Let θ_1 and θ_2 be suffix increasing functions wrt to their first argument.*

1. $\theta_1 \prec \theta_2$ *iff* $\mathbf{K}^{\theta_2}_{\text{spread}} <_{\text{ct}} \mathbf{K}^{\theta_1}_{\text{spread}}$
2. $\theta_1 \simeq \theta_2$ *iff* $\mathbf{K}^{\theta_2}_{\text{spread}} =_{\text{ct}} \mathbf{K}^{\theta_1}_{\text{spread}}$

The proof is similar to that of Theorem 43.

5 Discriminating modes

5.1 Universal modes

The notion of discriminating mode was introduced in Definition 21.

Lemma 68. *There is a recursive injective function $f : \Delta^* \to \Delta^*$ such that $W_{f(\mathtt{e})}$ is a maximal (with respect to inclusion) discriminating submode of $W_{\mathtt{e}}$.*

Proof. From $\mathtt{e}$ we (recursively) get a recursive increasing sequence $(R_n)_{n\in\mathbb{N}}$ of finite modes included in $W_{\mathtt{e}}$ such that the union of the R_n's is $W_{\mathtt{e}}$. We shall now write R for $W_{\mathtt{e}}$. Notice that the relation $(\mathtt{p}, x) \in R_n$ (in $\mathtt{p}, x, n$) is recursive. We then recursively define a sequence $(S_n)_n$ so that

(*) S_n is maximal among *discriminating* submodes of R_n which contain S_{n-1} (convention $S_{-1} = \emptyset$)

The existence of S_n is insured by the finiteness of R_n

We set S to be the union of the S_n's. Since the S_n's are increasing, we see that for every string x the sequence $(E_{S_n}(x))_{n\in\mathbb{N}}$ is (non strictly) decreasing with respect to the length-antilexicographic ordering on Δ^*. Thus, for every string x there exists k such that $E_{S_m}(x) = E_{S_n}(x)$ for all $m, n \geq k$. Hence, E_S is the pointwise limit of the E_{S_n}. Since the E_{S_n}'s are injective so is their limit E_S. Thus, S is a discriminating mode.

It remains to show that S is maximal among discriminating submodes of R. Take a discriminating mode T such that $S \subset T \subseteq R$. Suppose that $T \setminus S$ is not empty. Suppose that $(\mathtt{p}, x)$ is a pair of $T \setminus S$ for which $\mathtt{p}$ is least possible with respect to the length anti-lexicographic ordering on Δ^*. Suppose that $(\mathtt{p}, x) \in R_m$. There are two cases to consider.

1. $E_S(x) < \mathtt{p}$. Suppose $(E_S(x), x) \in S_n$ where $n \geq m$. Then $E_{S_n \cup \{(\mathtt{p},x)\}} = E_{S_n}$ so that $S_n \cup \{(\mathtt{p}, x)\}$ would also be a discriminating mode, contradicting the maximality of S_n.

2. Either $E_S(x) > \mathtt{p}$ or $E_S(x)$ is undefined. If $\mathtt{q} < \mathtt{p}$ then the very choice of $\mathtt{p}$ implies that $(\mathtt{q}, x) \notin T$. Thus, we have $E_T(x) = \mathtt{p}$. For each $(\mathtt{p}, y) \in S_m$, we have $(\mathtt{p}, y) \in T$, because $S_m \subseteq S \subset T$. Now, T is discriminating, so $E_T(y) < \mathtt{p}$. The minimality of $\mathtt{p}$ implies that $(E_T(y), y) \in S$. It follows that $S_m \cup \{(\mathtt{p}, x)\}$ is discriminating. The fact that S_m is maximal in R_m is violated.

Consequently, $S = T$, and it turns out that S is maximal. $\square$

Remark 69. Observe that the above construction is a greedy one.

Theorem 70. *There is a strong universal mode for discriminating modes. The entropy of discriminating modes is denoted* $\mathbf{K}_{\text{discri}}$.

Proof. Let $f : \Delta^* \to \Delta^*$ be the recursive injective function defined in Lemma 68. We define the mode U by $(\langle f(\mathtt{e}), \mathtt{p}\rangle, x) \in U$ holds iff $(\mathtt{p}, x) \in W_{f(\mathtt{e})}$. We now check that U is a discriminating mode. Suppose $E_U(x) = E_U(x')$ where $E_U(x) = \langle f(\mathtt{e}), \mathtt{p}\rangle$ and $E_U(x') = \langle f(\mathtt{e}'), \mathtt{p}'\rangle$. Then $\mathtt{e} = \mathtt{e}'$ and $\mathtt{p} = \mathtt{p}'$. Since $W_{f(\mathtt{e})}$ is discriminating, we conclude $x = x'$. It follows that U is a strong universal mode for discriminating modes. □

Remark 71. Clearly, U is a normal max-inclusive universal mode.

5.2 Relations between $\mathbf{K}_{\mathrm{det}}$ and $\mathbf{K}_{\mathrm{discri}}$

Lemma 72. *Let S be a discriminating mode, then*

$$\mathrm{card}\{x : (\mathtt{p}, x) \in S \textit{ and } |\mathtt{p}| \le n\} < 2^{n+1}$$

In particular, we see that a discriminating mode is a $\lambda n.n+1$-size mode.

Proof. Since $\{x : (x, \mathtt{p}) \in S$ and $|\mathtt{p}| \le n\} = \{x : |E_S(x)| \le n\}$ and E_S is injective, we have $\mathrm{card}\{x : |E_S(x)| \le n\} < 2^{n+1}$. So, the conclusion follows. □

Theorem 73. $\mathbf{K}_{\mathrm{det}} =_{\mathrm{ct}} \mathbf{K}_{\mathrm{discri}}$

Proof. By Lemma 72, a discriminating mode is a $\lambda n.(n+1)$-size mode. So, Theorem 41 yields $\mathbf{K}_{\mathrm{det}} \le_{\mathrm{ct}} \mathbf{K}_{\mathrm{discri}}$. Conversely, we have $\mathbf{K}_{\mathrm{discri}} \le_{\mathrm{ct}} \mathbf{K}_{\mathrm{det}}$ because any deterministic mode is a discriminating mode. □

6 Non-deterministic prefix modes

6.1 Prefix modes.

We begin by extending deterministic prefix modes to non-deterministic modes in the most obvious way.

Definition 74. Let R be a non deterministic mode.

1. R is *prefix* if the set of programs $\mathrm{dom}(R)$ is prefix free.
2. R is *Kraft* if $\sum_{\mathtt{p}\in\mathrm{dom}(R)} 2^{|\mathtt{p}|} \le 1$
3. R is *harmonic* if $\sum_{\mathtt{p}\in\mathrm{dom}(R)} 2^{|\mathtt{p}|}$ is convergent

It is easy to see that a prefix mode is a Kraft mode, and a Kraft mode is a harmonic mode.

Actually, as concerns entropies, there is no difference between description modes and prefix modes. Which is unlike the deterministic case.

Proposition 75. *For each description mode S, there is a prefix mode R such that $K_S =_{\mathrm{ct}} K_R$. Moreover, R is many-one reducible to S. A fortiori, the same is true with Kraft and harmonic modes.*

Proof. Put $f((\mathtt{p}, x)) = (1^{|p|}0, x)$. The function f is obviously recursive. Put $R = \{f((\mathtt{p}, x)) : (\mathtt{p}, x) \in S\}$. So, R is a weak prefix mode and is many-one reducible to S. Also, $K_R(x) = K_S(x) + 1$ for each x.

The other claims follow from the observed inclusions. □

6.2 Prefix-free set of minimal descriptions

Another direction is to require that the sole set of minimal descriptions (Definition 20) satisfies a prefix condition. For this purpose, define $\mathrm{range}(E_R) = \{\mathtt{p} : \exists x\, E_R(x) = \mathtt{p}\}$.

Definition 76.

1. A *min-prefix mode* R is a description mode R for which $\mathrm{range}(E_R)$ is prefix free.
2. A *min-Kraft mode* R is a description mode R for which
$$\sum_{\mathtt{p}\in\mathrm{range}(E_R)} 2^{|\mathtt{p}|} \le 1$$

3. A *min-harmonic mode* R is a description mode R for which

$$\sum_{\mathrm{p} \in \mathrm{range}(E_R)} 2^{|\mathrm{p}|} \text{ is convergent}$$

But again, the prefix mode R, which is constructed in the proof of Proposition 75 is, in fact, a min-prefix mode. Consequently, the class of modes thus defined does not delineate a new class of entropies.

6.3 Prefix-free set of minimal discriminating descriptions

Definition 77. A (non-deterministic) *discriminating min-prefix mode* is description mode R which is both min-prefix and discriminating, i.e. such that

1. $\mathrm{range}(E_R)$ is prefix free,
2. E_R is an injective function.

The two other prefix-like conditions readily adapt as follows.

Definition 78.

1. A *discriminating min-Kraft mode* R is a min-Kraft mode for which E_R is injective.
2. A *discriminating min-harmonic mode* R is a min-harmonic mode for which E_R is injective.

Theorem 79. *1. There is a strong universal mode for discriminating min-Kraft modes. The associated entropy is denoted* $\mathbf{K}_{\mathrm{dis.minkraft}}$

2. There is a strong universal mode for discriminating min-harmonic modes. The associated entropy is denoted $\mathbf{K}_{\mathrm{dis.minharmonic}}$

Proof. The proof of (1) is a straightforward adaptation of Theorem 25 in order to enumerate min-Kraft modes and of Lemma 68 in order to construct maximal discriminating modes.

The proof of (2) goes as follows. Say that a discriminating min-harmonic mode is c-min-harmonic if it satisfies $\sum_{x\in\Sigma^*} 2^{|E_R(x)|} \leq c$. As in the previous case, there is a strong universal mode for c-min-harmonic modes. To obtain an universal mode for discriminating min-harmonic modes, we follow Theorem 70. The sole modification is that the function f has now two parameters (the index e and a constant c) and computes an index for $W_{f(\mathsf{e},c)}$. Moreover, $W_{f(\mathsf{e},c)}$ is a maximal mode included in W_e. □

Remark 80. The second parameter c in the construction of the min-harmonic universal mode prevents this mode to be effective universal (cf. Definition 9). In fact, one can prove that there is no effective universal mode in the class of min-harmonic modes. The reason is as follows :

1. $\{\mathsf{e} : W_\mathsf{e} = U_e\}$ is Π^0_2,
2. $\{\mathsf{e} : W_\mathsf{e}$ is min-harmonic$\}$ is Σ^0_2-complete,
3. if U were effectively universal then the two above sets would be equal.

Except for the harmonic and min-harmonic classes, all universal modes defined in this paper are effective.

In order to get a discriminating min-prefix mode we first prove the analog of Lemma 68.

Lemma 81. *There is a recursive injective function $f : \Delta^* \to \Delta^*$ such that*

1. *$W_{f(\mathsf{e})}$ is a discriminating min-prefix mode included in W_e,*
2. *$W_{f(\mathsf{e})} = W_\mathsf{e}$ whenever W_e is itself discriminating min-prefix.*

Proof. From e we (recursively) get a recursive increasing sequence $(R_n)_n$ of finite modes included in W_e such that the union of the $(R_n)_n$'s is W_e. We shall now write R for W_e. Notice that the relation $(\mathsf{p}, x) \in R_n$ is recursive.

We then recursively define a sequence (S_n) of finite discriminating min-prefix modes as follows

1. $S_0 = \emptyset$

2. If $E_{R_{n+1}} = E_{S_n}$ then (R_{n+1} is surely discriminating min-prefix as is S_n and) we let $S_{n+1} = R_{n+1}$

3. Else, let x be the smallest string such that

   either $E_{R_{n+1}}(x) < E_{S_n}(x)$ or $E_{R_{n+1}}(x)$ is defined but not $E_{S_n}(x)$

   If there is some discriminating min-prefix submode M of R_{n+1} which contains $S_n \cup \{(E_{R_{n+1}}(x), x)\}$, then consider a maximal (wrt inclusion) such M and set $S_{n+1} = M$.

   Otherwise, we freeze the construction by setting $S_{n+1} = S_n$.

We set S to be the union of the S_n's. Observe that the function E_S is the pointwise limit of the E_{S_n}'s. Since the S_n's are discriminating min-prefix so is S.

Observe the following property :

(*) If $\mathtt{p} \geq E_S(x)$ and $(\mathtt{p}, x) \in R$ then $(\mathtt{p}, x) \in S$

In fact, if $(E_S(x), x) \in S_n$ then $E_{S_n}(x) = E_S(x)$ so that adding any $(\mathtt{p}, x) \in R_{n+1}$ to any supermode of S_n can not destroy its discriminating min-prefix character. Thus, the maximality of M implies that $(\mathtt{p}, x) \in S_{n+1}$.

It remains to show that if R is discriminating min-prefix, then $R = S$. By way of contradiction, suppose that $R \setminus S$ is non empty. Using (*) there must be some x such that $E_R(x) < E_S(x)$. Consider the least such x. Then

$$\forall y < x\, (E_R(y) \,\text{is defined} \;\Rightarrow (E_R(y) = E_S(y) \wedge (E_S(y), y) \in S))$$

Let m be such that

$$(E_R(x), x) \in R_m \text{ and } \forall y < x\, (E_R(y) \,\text{is defined} \;\Rightarrow (E_R(y), y) \in S_m)).$$

Then x is the smallest string considered in Item 3 of the construction of S_n for every $n \geq m$.

By very construction of the S_n's, we see that after step m they remain stationary : $S_m = S_{m+1} = \ldots = S$.

However, $\text{range}(R_m) = \{x \ : \ \exists \mathtt{p}\, (\mathtt{p}, x) \in R\}$ is finite so that there is some $n > m$ such that $(E_R(y), y) \in R_n$ for all $y \in \text{range}(R_m)$. Hence $R_n \cap (\Delta^* \times \text{range}(R_m))$ is discriminating min-prefix. Since it properly contains S_n (which is equal to S_m), this contradicts Item 3 of the above construction. □

Remark 82. The discriminating min-prefix submode S of R can not be taken maximal in general. This is in contrast with Lemma 68. The reason is that a greedy algorithm to get a maximal submode does not work. This is illustrated by the following example. Let R be the discriminating min-prefix enumerated as follows.

$$
\begin{aligned}
&R_0 = \emptyset \\
&R_1 = R_0 \cup \{(\mathtt{11}, 0), (\mathtt{011}, 1), (\mathtt{0011}, 2)\} \\
&R_2 = R_1 \cup \{(\mathtt{01}, 0), (\mathtt{00011}, 3)\} \\
&R_3 = R_2 \cup \{(\mathtt{001}, 1), (\mathtt{000011}, 4)\} \\
&R_4 = R_3 \cup \{(\mathtt{0001}, 2), (\mathtt{0000011}, 5)\} \\
&\ldots
\end{aligned}
$$

Then, a greedy algorithm would compute

$$
\begin{aligned}
&S_0 = \emptyset \\
&S_1 = R_1 \\
&S_2 = S_1 \cup \{(\mathtt{00011}, 3)\} \\
&S_3 = S_2 \cup \{(\mathtt{000011}, 4)\} \\
&S_4 = S_2 \cup \{(\mathtt{0000011}, 5)\} \\
&\ldots
\end{aligned}
$$

Indeed, it was not possible to put $(\mathtt{01}, 0)$ in S_2 (and preserve the min-prefix character) because $(\mathtt{011}, 1)$ had already been put in S_1. Similarly with other steps. Thus, though R is discriminating min-prefix, S is strictly included in R.

Theorem 83. *There is a strong universal mode for discriminating min-prefix modes. The entropy of discriminating min-prefix modes is denoted* $\mathbf{K}_{\text{dis.minprefix}}$.

Proof. As in the proof of discriminating modes, let $f : \Delta^* \to \Delta^*$ be as in Lemma 81. Put $U = \{(\langle f(\mathtt{e}), \mathtt{p}\rangle, x) : (\mathtt{p}, x) \in W_{f(\mathtt{e})}\}$. We see that the $U_{\mathtt{e}}$'s constitute an enumeration of discriminating min-prefix modes. We prove that U is discriminating exactly as in Theorem 70.

Now observe that $\langle f(\mathtt{e}), \mathtt{p}\rangle$ is a prefix of $\langle f(\mathtt{e}'), \mathtt{q}\rangle$ only if $\mathtt{e} = \mathtt{e}'$ and $\mathtt{p}$ is a prefix of $\mathtt{p}'$. Since each $W_{f(\mathtt{e})}$ is min-prefix, so is U. □

Remark 84. The discriminating min-prefix universal mode defined above is not max-inclusive, but is effective.

The next Proposition, due to Levin [6] and also reported in [11] claims that for each non-deterministic prefix mode, there is a deterministic prefix mode with the same set of minimal descriptions and consequently the same entropy.

Definition 85. A function $f : \Delta^* \to \mathbb{N}$ is recursively enumerable from above if the set $\{(\mathtt{p}, n) : f(\mathtt{p}) \leq n\}$ is recursively enumerable.

Proposition 86. *Let f be a recursively enumerable function from above such that $\sum_{\mathtt{p}\in\Delta^*} 2^{-f(\mathtt{p})}$ is convergent. Then, there is a deterministic prefix mode S such that $K_S =_{\text{ct}} f$.*

Proposition above yields that for each discriminating min-harmonic mode R there is a deterministic prefix mode S such that $K_R =_{\text{ct}} K_S$. In particular, this also holds for discriminating min-kraft and min-prefix modes.

Theorem 87.

$$\mathbf{KP} =_{\mathrm{ct}} \mathbf{K}_{\mathrm{dis.minprefix}} =_{\mathrm{ct}} \mathbf{K}_{\mathrm{dis.minkraft}} =_{\mathrm{ct}} \mathbf{K}_{\mathrm{dis.minharmonic}}$$

Remark 88. We can say that the above results on min-prefix modes show the robustness of the notion of deterministic description modes. The fundamental reason lies on a result of Levin in [4, 5, 6] where it is demonstrated that the series $\sum_{x\in\Sigma^*} 2^{-\mathbf{KP}(x)}$ is the greatest real valued r.e. series which converges, up to a multiplicative factor. (By a real valued r.e. series f, we mean that f is recursively enumerable from below (when considering rational Dedekind cuts), i.e. $\{y \in \mathbb{Q} : y \leq f(x)\}$ is recursively enumerable).

References

[1] G. Chaitin. A theory of program size formally identical to information theory. *JACM*, 22:329–340, 1975.

[2] N. Jones. *Computability and complexity, from a programming perspective.* MIT press, 1997.

[3] L. Levin. Laws of information conservation (non-growth) and aspects of the foundation of probability theory. *Problems Inform. Transmission*, 10(3):206–210, 1974.

[4] L. Levin. Uniform tests of randomness. *Soviet. Math. Dokl.*, 17(2):337–340, 1976.

[5] L. Levin. Various measure of complexity for finite objects (axiomatic description). *Soviet. Math. Dokl.*, 17(2):522–526, 1976.

[6] L. Levin. Randomness conservation inequalities: information and independence in mathematical theories. *Information and Computation*, 61:15–37, 1984.

[7] M. Li and P. Vitányi. *An Introduction to Kolmogorov Complexity and its Application.* Springer, 1997. (Second edition).

[8] D. Loveland. A variant of the Kolmogorov concept of complexity. *Information and control*, (15):510–526, 1969.

[9] P. Orponen, K. Ko, U. Schöning, and O. Watanabe. Instance complexity. *JACM*, pages 96–121, 1994.

[10] V. Uspensky. Kolmogorov and mathematical logic. *Journal of Symbolic Logic*, 57(2):385–412, June 1992.

[11] V. Uspensky and A. Shen. Relations between varieties of Kolmogorov complexities. *Mathematical systems theory*, 29:271–292, 1996.