DIMACS Technical Report 97-74

Upper semi-lattice of binary strings with the relation "x is simple conditional to y"

by

Andrei Muchnik
Institute of New Technologies, 10 Nizhnyaya Radischewskaya
Moscow, Russia 109004
e-mail: amuchnik@int.glasnet.ru

Andrei Romashchenko Nikolai Vereshagin[1]
Dept. of Mathematical Logic and Theory of Algorithms
Moscow State University
Vorobjewy Gory, Moscow, Russia 119899
email: an@romash.mccme.ru ver@mech.math.msu.su

Alexander Shen
Institute of Problems of Information Transmission
email: shen@mccme.ru

[1]The work was partially done while visiting the University of Amsterdam and DIMACS center at Rutgers University

---

DIMACS is a partnership of Rutgers University, Princeton University, AT&T Labs-Research, Bell Labs and Bellcore.

DIMACS is an NSF Science and Technology Center, funded under contract STC–91–19999; and also receives support from the New Jersey Commission on Science and Technology.

ABSTRACT

We study the properties of the set of binary strings with the relation "the Kolmogorov complexity of x conditional to y is small". We prove that there are pairs of strings which have no greatest common lower bound with respect to this pre-order. We present several examples when the greatest common lower bound exists but its complexity is much less than mutual information (extending Gács and Körner result [2]).

1 Introduction

The family of Turing degrees is a well studied object. Is there any finite analog of it? In the present paper we study such a finite analog: instead of subsets of $\mathbb{N}$ we take binary strings and instead of Turing reducibility we take the relation "Kolmogorov complexity of x conditional to y is small". This structure has many properties common with the set of Turing degrees. For example, it forms also an upper semi-lattice. And it is more rich, since we can measure the complexity of finite strings.

Of course, we should make precise what means that the Kolmogorov complexity of x conditional to y, $K(x|y)$, is small. Instead of being small we will speak about being less than $c\log(|x|+|y|)$ [1], where $|\cdot|$ denotes the length. So we say that x is c-simple conditional to y, $x \leq_c y$, if $K(x|y) < c(\log K(x) + \log K(y))$. Then the above statement that the structure $\langle\{0,1\}^*, \leq_c\rangle$ is an upper semi-lattice means the following:
1) the existence of the lowest common upper bound:

$$\exists c_1\ \forall c_2\ \exists c_3\ \forall xy\ \exists z\ (x, y \leq_{c_1} z\ \wedge\ \forall z'\ (x, y \leq_{c_2} z' \Rightarrow z \leq_{c_3} z'));$$

2) the non-existence of the greatest common lower bound:

$$\forall c_1 \exists c_2\ \forall c_3\ \exists xy\ \forall z\ (x, y \geq_{c_1} z \Rightarrow \exists z'\ (x, y \geq_{c_2} z'\ \wedge\ z \not\geq_{c_3} z')).$$

Why not to define that x is c-simple conditional to y by inequality $K(x|y) \leq c$? The reason is that many interesting inequalities for Kolmogorov complexity are true only up to $\mathcal{O}(\log n)$ term, where n is the sum of the complexities of all strings involved; for instance the well known commutativeness law for information states that $I(x:y) - I(y:x) = \mathcal{O}(\log K(x) + \log K(y))$ [2] Another reason is that different versions of Kolmogorov complexity are equivalent up to logarithmic term (see [9]). Therefore the relation $K(x|y) \leq c(\log K(x) + \log K(y))$ is more interesting than the relation $K(x|y) \leq c$.

Another, more simple, way to make the results rigorous is to consider infinite sequences of binary strings rather then single strings. We say that a sequence $\boldsymbol{x} = \{x_n\}$ is simple conditional to a sequence $\boldsymbol{y} = \{y_n\}$, $\boldsymbol{x} \leq \boldsymbol{y}$, if $K(x_n|y_n) = \mathcal{O}(\log n)$. We restrict ourselves with sequences $\boldsymbol{x}$ satisfying the requirement $|x_n| \leq \mathrm{poly}(n)$, where $\mathrm{poly}(n)$ is an arbitrary polynomial. The set of all such sequences is denoted by R. Using the pre-ordered set $\langle R, \leq\rangle$ we can restate two above assertions 1) and 2) using the standard notions of greatest common lower bound and lowest common upper bound in a pre-ordered set:
1') the existence of the lowest common upper bound:

$$\forall \boldsymbol{x}\boldsymbol{y}\ \exists \boldsymbol{z}\ (\boldsymbol{x}, \boldsymbol{y} \leq \boldsymbol{z}\ \wedge\ \forall \boldsymbol{z}'\ (\boldsymbol{x}, \boldsymbol{y} \leq \boldsymbol{z}' \Rightarrow \boldsymbol{z} \leq \boldsymbol{z}'));$$

2') the non-existence of the greatest common lower bound:

$$\exists \boldsymbol{x}\boldsymbol{y}\ \forall \boldsymbol{z}\ (\boldsymbol{x}, \boldsymbol{y} \geq \boldsymbol{z} \Rightarrow \exists \boldsymbol{z}'\ (\boldsymbol{x}, \boldsymbol{y} \geq \boldsymbol{z}'\ \wedge\ \boldsymbol{z} \not\geq \boldsymbol{z}')).$$

[1] The base of all logarithms in the paper is 2. Moreover, for technical reasons, we let $\log x = \log_2(x+2)$, since we want $\log n$ to be positive for all integer n

[2] $f(n) = \mathcal{O}(g(n))$ means that $|f(n)| \leq C|g(n)|$ for some C for all n. So $\mathcal{O}(g(n))$ may be negative.

In this form, the results seems to be a bit weaker, but their formulations become simpler. Therefore, in this paper, we prefer to formulate our results using the structure $\langle R, \leq\rangle$ or the structure $\langle R, \leq, K(\cdot)\rangle$, where $K(\cdot)$ is the Kolmogorov complexity.

So, our first result is the assertion 2' (and the assertion 1' is easy).

Our next results deal with common information. Gács and Körner were first who pointed out that the quantity of mutual information $I(x:y)$ may be large when the strings x and y have nothing in common. More specifically, Gács and Körner in [2] proved that there exist x_n and y_n of length n such that the mutual information $I(x_n : y_n)$ has linear growth cn and the complexity of any common lower bound of x_n and y_n is $\mathcal{O}(\sqrt{n})$ (Another proof appeared in [6].) Using another approach, we present three examples of sequences x_n and y_n of length n such that $I(x_n : y_n)$ has linear growth and the complexity of any common lower bound of $\boldsymbol{x}$ and $\boldsymbol{y}$ is $\mathcal{O}(\log n)$. Our proof of this fact is simpler than the rather technical proof of the corresponding fact in [2].

If two sequences have greatest common lower bound, its complexity, defined up to $\mathcal{O}(\log n)$ term, is called *common information* in those sequences, as opposed to mutual information $I(x_n : y_n)$. It is well known that any common lower bound $\boldsymbol{z}$ of sequences $\boldsymbol{x}$ and $\boldsymbol{y}$ satisfies the inequality $K(z_n) \leq I(x_n : y_n) + \mathcal{O}(\log n)$ (for the seek of completeness, this fact is proven in the next section), thus the common information does not exceed the mutual information. The amount of common information does not determine completely how much $\boldsymbol{x}$ and $\boldsymbol{y}$ have in common. What reflects this better is the set $M_{\boldsymbol{xy}}$ of all triples (u, v, w) of reals such that there exists $\boldsymbol{z}$ for which

$$K(z_n) \leq un + \mathcal{O}(\log n),\ K(x_n|z_n) \leq vn + \mathcal{O}(\log n),\ K(y_n|z_n) \leq wn + \mathcal{O}(\log n).$$

(Here, we assume that $K(x_n), K(y_n), I(x_n : y_n)$ have linear growth. Moreover, we we are mostly interested in the specific case when $K(x_n) = 2n + \mathcal{O}(\log n)$, $K(y_n) = 2n + \mathcal{O}(\log n)$, $K(\langle x_n, y_n\rangle) = 3n + \mathcal{O}(\log n)$.) The larger is this set the more common have $\boldsymbol{x}, \boldsymbol{y}$. We prove (in a sense non-constructively) that there exist $\boldsymbol{x}, \boldsymbol{y}$ with given complexities $K(x_n)$, $K(y_n)$, $I(x_n : y_n)$ having minimum possible set $M_{\boldsymbol{xy}}$. (Our proof is a generalization of the proof in [6, 7].) We study then what is the set $M_{\boldsymbol{xy}}$ for one interesting particular pair $(\boldsymbol{x}, \boldsymbol{y})$.

The paper is organized as follows. In Section 2, we recall basic definitions. In Section 3, we prove that there are sequences having no greatest common lower bound. In Section 4, we present the examples of sequences in which common information is less than mutual information (including Gács and Körner's example). In Section 5, we study the set $M_{\boldsymbol{xy}}$ for certain $\boldsymbol{x}, \boldsymbol{y}$. And in Section 6, we establish some simple properties of the structure $\langle R, \leq\rangle$.

2 Preliminaries

The set of finite binary strings is denoted by Ξ, the empty string is denoted by ϵ. A *conditional description method* is a partial computable function $\phi : \Xi \times \Xi \to \Xi$. The string p is called a *program* to compute x given y by ϕ if

$$\phi(p, y) = x.$$

The complexity $K_\phi(x|y)$ of x conditional to y with respect to ϕ is defined by

$$K_\phi(x|y) = \min\{|p| \mid \phi(p, y) = x\}$$

and $K_\phi(x|y) = \infty$ if there is no such p. An invariance theorem by Solomonoff, Kolmogorov and Chaitin [8, 3, 1] states that there exists *universal conditional description method,* that is, ϕ_0 such that for any ϕ there is a constant c_ϕ for which

$$K_{\phi_0}(x|y) \leq K_\phi(x|y) + c_\phi$$

for all x, y. This theorem is easy to show. Indeed, let

$$\phi_0(p, y) = x$$

iff p has the form $p_1 p_2$, where p_1 is a program to compute a function $\phi : \Xi \times \Xi \to \Xi$ such that $\phi(p_2, y) = x$ (we assume that, in the programming language, no program p_1 is a prefix of another program p_1', thus it is possible to find p_1, p_2 given their concatenation $p_1 p_2$). If $\phi(p_2, y) = x$ and p_1 is a program to compute ϕ, then $\phi_0(p_1 p_2, y) = x$, therefore

$$K_{\phi_0}(x|y) \leq K_\phi(x|y) + |p_1|$$

for all x, y. Hence ϕ_0 is universal.

We fix a universal conditional description method ϕ_0 and define *complexity of x conditional to y* by

$$K(x|y) = K_{\phi_0}(x|y)$$

and *complexity of x* by

$$K(x) = K(x|\epsilon).$$

If $\phi_0(p, \epsilon) = x$ we say that p is a program to compute x.

Let us fix a computable one to one correspondence between $\Xi \times \Xi$ and Ξ and denote by $\langle u, v\rangle$ the binary string corresponding to the pair (u, v).

It is easy to show that

$$K(\langle x, y\rangle) \leq K(y) + K(x|y) + 2 \log K(y) + \mathcal{O}(1).$$

Indeed, let p be a minimal program to compute y and q be a minimal program to compute x given y. Let us double all the bits in the binary representation of the length of p and add the string 01 to the end, denote the resulting string by $l(p)$ (for example, if p has 5 bits, then $l(p) = 11001101$). Given the string $l(p)pq$ we can find both p and q, then find x and y, and finally find $\langle x, y\rangle$.

It turns out that vice versa

$$K(y) + K(x|y) \leq K(\langle x, y\rangle) + \mathcal{O}(\log K(x) + \log K(y))$$

(see [5], or [10]), thus

$$K(y) + K(x|y) = K(\langle x, y\rangle) + \mathcal{O}(\log K(x) + \log K(y)). \tag{1}$$

We will use this equation several times.

The mutual information in x and y is defined by

$$I(x:y) = K(y) - K(y|x).$$

The equality (1) implies that the mutual information is commutative up to a logarithmic term:

$$\begin{aligned} I(x:y) &= I(y:x) + \mathcal{O}(\log K(x) + \log K(y)) \\ &= K(x) + K(y) - K(\langle x, y\rangle) + \mathcal{O}(\log K(x) + \log K(y)). \end{aligned}$$

It is easy to see that the number of binary strings of length less than n is equal to $1 + 2 + \cdots + 2^{n-1} = 2^n - 1$. Therefore any set A of strings having at least 2^n elements has a string x with $K(x) \geq n$. Any such string is called a *random* string in A. So, for instance, a random string of length n is a string of length n and complexity at least n. Similarly, for any finite A, B there is a pair (a, b) in $A \times B$ such that $K(\langle a, b\rangle) \geq [\log(|A||B|)]$. Any such a, b are called *random independent* elements of A and B, respectively.

3 The structure $\langle R, \leq, K(\cdot)\rangle$

Recall that we denote by R the set of all sequences $\boldsymbol{x}$ of binary strings such that the length of x_n is bounded by a polynomial of n. By Λ we denote the sequence, whose all terms are equal to the empty string.

Definition 1 *A sequence $\boldsymbol{x}$ is simple conditional to a sequence $\boldsymbol{y}$, $\boldsymbol{x} \leq \boldsymbol{y}$, if*

$$K(x_n|y_n) = \mathcal{O}(\log n).$$

A sequence $\boldsymbol{z}$ is called a common lower [upper] bound of sequences $\boldsymbol{x}$ and $\boldsymbol{y}$ if $\boldsymbol{z} \leq \boldsymbol{x}$ and $\boldsymbol{z} \leq \boldsymbol{y}$ [$\boldsymbol{z} \geq \boldsymbol{x}$ and $\boldsymbol{z} \geq \boldsymbol{y}$]. A sequence $\boldsymbol{z} \in R$ is called a lowest common upper bound of sequences $\boldsymbol{x}, \boldsymbol{y} \in R$, $\boldsymbol{z} = \sup(\boldsymbol{x}, \boldsymbol{y})$, if $\boldsymbol{x} \leq \boldsymbol{z}$, $\boldsymbol{y} \leq \boldsymbol{z}$ and if a sequence $\boldsymbol{z}'$ satisfies $\boldsymbol{x} \leq \boldsymbol{z}'$, $\boldsymbol{y} \leq \boldsymbol{z}'$ then $\boldsymbol{z} \leq \boldsymbol{z}'$. The notion of the greatest common lower bound, $\inf(\boldsymbol{x}, \boldsymbol{y})$, is defined analogously. Obviously, if $\sup(\boldsymbol{x}, \boldsymbol{y})$ [$\inf(\boldsymbol{x}, \boldsymbol{y})$] exists it is not unique, though it is unique up to equivalence relation $\boldsymbol{z} \leq \boldsymbol{z}' \wedge \boldsymbol{z}' \leq \boldsymbol{z}$.

Proposition 1 *Any two elements $\boldsymbol{x}$, $\boldsymbol{y}$ in R have lowest common upper bound.*

Proof. Let $z_n = \langle x_n, y_n\rangle$. $\square$

In the following proposition we establish the upper bound for the complexity of any common lower bound of $\boldsymbol{x}$ and $\boldsymbol{y}$.

Proposition 2 *If $\boldsymbol{z} \leq \boldsymbol{x}$ and $\boldsymbol{z} \leq \boldsymbol{y}$, then $K(z_n) \leq I(x_n : y_n) + \mathcal{O}(\log n)$.*

Proof. We have

$$\begin{aligned} K(\langle x_n, z_n\rangle) &= K(x_n) + K(z_n|x_n) + \mathcal{O}(\log n) = K(x_n) + \mathcal{O}(\log n) \\ K(x_n|y_n) &\leq K(x_n|z_n) + K(z_n|y_n) + \mathcal{O}(\log n) = K(x_n|z_n) + \mathcal{O}(\log n) \\ I(x_n : y_n) &= K(x_n) - K(x_n|y_n) \geq K(\langle x_n, z_n\rangle) - K(x_n|z_n) + \mathcal{O}(\log n) \\ &= K(z_n) + \mathcal{O}(\log n).\Box \end{aligned}$$

If $\boldsymbol{z} \leq \boldsymbol{x}$, $\boldsymbol{z} \leq \boldsymbol{y}$ and $K(z_n) = I(x_n : y_n) + \mathcal{O}(\log n)$, we say that $\boldsymbol{z}$ *extracts the mutual information from* $\boldsymbol{x}$ *and* $\boldsymbol{y}$.

Proposition 3 *If* $\boldsymbol{z}$ *extracts the mutual information from* $\boldsymbol{x}$ *and* $\boldsymbol{y}$*, then* $\boldsymbol{z}$ *is the greatest common lower bound of* $\boldsymbol{x}$ *and* $\boldsymbol{y}$.

Proof. Let $\boldsymbol{u} \leq \boldsymbol{x}$ and $\boldsymbol{u} \leq \boldsymbol{y}$. The sequence $\langle z_n, u_n\rangle$ is a common lower bound of $\boldsymbol{x}$ and $\boldsymbol{y}$. Using Proposition 2 we get $K(\langle z_n, u_n\rangle) \leq I(x_n : y_n) + \mathcal{O}(\log n) = K(z_n) + \mathcal{O}(\log n)$. Hence $K(u_n|z_n) = K(\langle z_n, u_n\rangle) - K(z_n) + \mathcal{O}(\log n) \leq \mathcal{O}(\log n)$. $\Box$

If $\boldsymbol{x}$ and $\boldsymbol{y}$ do have the greatest common lower bound $\boldsymbol{z}$, we call, after Gács and Körner [2], the complexity of $\boldsymbol{z}$ the *common information in* $\boldsymbol{x}$ and $\boldsymbol{y}$. The common information is defined up to logarithmic term.

Many pairs of sequences have the lowest common upper bound. Let us give an example of such a pair.

Example 1 *Let* $x_n = \langle p_n, q_n\rangle$*,* $y_n = \langle p_n, r_n\rangle$*, where* $|p_n| = |q_n| = |r_n| = n$ *and* $p_n q_n r_n$ *is a random string of length* $3n$*. Then* $\boldsymbol{p} = \{p_n\}$ *is the greatest common lower bound of* $\boldsymbol{x}$ *and* $\boldsymbol{y}$ *and extracts the mutual information from* $\boldsymbol{x}, \boldsymbol{y}$*. Indeed, we have* $K(x_n) = 2n + \mathcal{O}(\log n)$*,* $K(y_n) = 2n + \mathcal{O}(\log n)$*,* $K(\langle x_n, y_n\rangle) = 3n + \mathcal{O}(\log n)$*,* $I(x_n : y_n) = n + \mathcal{O}(\log n)$ *and* $K(p_n) = n + \mathcal{O}(\log n)$*. And* $\boldsymbol{p} \leq \boldsymbol{x}$ *and* $\boldsymbol{p} \leq \boldsymbol{y}$*, thus* p_n *extracts the mutual information from* $\langle p_n, q_n\rangle$ *and* $\langle p_n, r_n\rangle$*. By Proposition 3 the sequence* $\boldsymbol{p}$ *is the greatest common lower bound of* $\boldsymbol{x}$ *and* $\boldsymbol{y}$.

Theorem 1 *There are* $\boldsymbol{x}, \boldsymbol{y} \in R$ *having no greatest common lower bound.*

Proof. Let w_n be a random binary string of length $2n^2 + n$. Split w_n into $2n + 1$ blocks of length n: $w_n = \varepsilon_n b_1^0 b_2^0 \cdots b_n^0 b_1^1 b_2^1 \cdots b_n^1$. Let $\varepsilon_n^1, \varepsilon_n^2, \ldots, \varepsilon_n^n$ be the bits of ε_n. Let $x_n = b_1^0 b_2^0 \cdots b_n^0 b_1^1 b_2^1 \cdots b_n^1$ and $y_n = b_1^{\varepsilon_n^1} b_2^{\varepsilon_n^2} \cdots b_n^{\varepsilon_n^n}$.

To prove that $\boldsymbol{x}$ and $\boldsymbol{y}$ have no greatest common lower bound assume, for the contrary, that $\boldsymbol{z} = \inf(\boldsymbol{x}, \boldsymbol{y})$. Let us prove that there exists C such that $K(b_k^{\varepsilon_n^k}|z_n) \leq C \log n$ for all n and all $k \leq n$.

For every n let $k = k_n$ be the index such that the complexity $K(b_k^{\varepsilon_n^k}|z_n)$ is maximal among complexities $K(b_{k'}^{\varepsilon_n^{k'}}|z_n)$. Let $u_n = b_{k_n}^{\varepsilon_n^{k_n}}$. Evidently, $\boldsymbol{u}$ is a common lower bound of $\boldsymbol{x}$ and $\boldsymbol{y}$. Therefore, $\boldsymbol{u} \leq \boldsymbol{z}$, i.e. there is a C such that $K(b_k^{\varepsilon_n^{k_n}}|z_n) \leq C \log n$ for all n. Hence,

$$K(b_{k_n}^{\varepsilon_n^k}|z_n) \leq C \log n \tag{2}$$

for all k, n.

Lemma 1 *For large enough n, for all $i \in \{0,1\}$, $k \leq n$ the following holds:*

$$K(b_k^i|z_n) \leq C\log n \Longleftrightarrow i = \varepsilon_n^k.$$

Proof. The implication from the left to the right holds by (2).

To prove the inverse, assume that for some $r \leq n$, $K(b_r^{\bar{\varepsilon}_n^r}|z_n) \leq C\log n$, where $\bar{\varepsilon}_n^r = 1-\varepsilon_n^r$. Then $K(b_r^{\bar{\varepsilon}_n^r}|y_n) \leq K(z_n|y_n) + \mathcal{O}(\log n) = \mathcal{O}(\log n)$.

Let us note that w_n can be retrieved from y_n, ε_n and $b_1^{\bar{\varepsilon}_n^1}\cdots b_n^{\bar{\varepsilon}_n^n}$. And $b_r^{\bar{\varepsilon}_n^r}$ is simple conditional to y_n. Therefore w_n can be retrieved from y_n, ε_n, $b_1^{\bar{\varepsilon}_n^1}\cdots b_{r-1}^{\bar{\varepsilon}_n^{r-1}} b_{r+1}^{\bar{\varepsilon}_n^{r+1}}\cdots b_n^{\bar{\varepsilon}_n^n}$ and logarithmic number of extra bits.

Thus,

$$\begin{aligned}
K(w_n) &\leq K(y_n) + K(w_n|y_n) + \mathcal{O}(\log n) \leq \\
&\leq K(y_n) + K(\varepsilon_n) + K(b_1^{\bar{\varepsilon}_n^1}\cdots b_{r-1}^{\bar{\varepsilon}_n^{r-1}} b_{r+1}^{\bar{\varepsilon}_n^{r+1}}\cdots b_n^{\bar{\varepsilon}_n^n}) + \mathcal{O}(\log n) \\
&\leq n^2 + n + n(n-1) + \mathcal{O}(\log n) \\
&\leq 2n^2 + \mathcal{O}(\log n).
\end{aligned}$$

As $K(w_n) \geq 2n^2 + n$, for large enough n there is no r such that $K(b_r^{\bar{\varepsilon}_n^r}|z_n) \leq C\log n$. □

The above lemma yields a way to obtain ε_n from x_n using $\mathcal{O}(\log n)$ help bits: find first z_n, then to find ε_n^k find the i such that $K(b_k^i|z_n) \leq C\log n$. Hence, $K(\varepsilon_n|x_n) \leq \mathcal{O}(\log n)$ and $K(w_n) = K(x_n) + K(\varepsilon_n|x_n) + \mathcal{O}(\log n) = n^2 + \mathcal{O}(\log n)$. The obtained contradiction proves the theorem. □

Corollary 2 ([2, 6]) *There are sequences $\boldsymbol{x}, \boldsymbol{y}$ having no sequences extracting their mutual information.*

4 Examples of pairs of sequences whose common information is less than their mutual information

4.1 Gács and Körner's example

Let α, β be binary random values. Consider n outcomes of independent trials $\langle\alpha_1, \beta_1\rangle$, $\langle\alpha_2, \beta_2\rangle$, …, $\langle\alpha_n, \beta_n\rangle$ of $\langle\alpha, \beta\rangle$ and let $x_n = \alpha_1\alpha_2\ldots\alpha_n$ and $y_n = \beta_1\beta_2\ldots\beta_n$. It is easy to show that with high probability $I(x_n : y_n) = J(\alpha : \beta)n + o(n)$, where $J(\alpha : \beta)$ is the Shannon information of α, β. If α and β are dependent, then $J(\alpha : \beta) > 0$ and $I(x_n : y_n)$ has linear growth. Gács and Körner proved that if no more than one of the values

$$\begin{array}{cc}
\mathrm{Prob}[\alpha = \beta = 1], & \mathrm{Prob}[\alpha = \beta = 0], \\
\mathrm{Prob}[\alpha = 0, \beta = 1], & \mathrm{Prob}[\alpha = 1, \beta = 0]
\end{array}$$

is zero, then the complexity of any common lower bound of $\boldsymbol{x}, \boldsymbol{y}$ is $\mathcal{O}(\sqrt{n})$. So let for example

$$\begin{aligned} \mathrm{Prob}[\alpha = \beta = 1] = 0.4, &\qquad \mathrm{Prob}[\alpha = \beta = 0] = 0.4, \\ \mathrm{Prob}[\alpha = 0, \beta = 1] = 0.1, &\qquad \mathrm{Prob}[\alpha = 1, \beta = 0] = 0.1. \end{aligned}$$

Then the common information in $\boldsymbol{x}$ and $\boldsymbol{y}$ is $\mathcal{O}(\sqrt{n})$ but the mutual information has linear growth.[3]

4.2 Orthogonal lines in $\mathbb{F}_q^3$

Let us fix for every integer n an exponent of a prime $q = q_n$ such that $[\log q] = n$. So n is about the number of bits in binary representation of q.

A line in $\mathbb{F}_q^3$ is a 1-dimensional subspace of the linear space $\mathbb{F}_q^3$. There are $q^2 + q + 1$ different lines. Let l, l' be two lines with leading vectors (a, b, c) and (a', b', c'). They are orthogonal if $aa' + bb' + cc' = 0$. Let U be the set of all pairs of lines (x, y) such that x and y are orthogonal.

There are $q + 1$ lines orthogonal to any line, therefore $|U| = (q^2 + q + 1)(q + 1)$. Let (x_n, y_n) be a random pair in U. It is easy to see that $K(\langle x_n, y_n \rangle) = 3n + \mathcal{O}(\log n)$, $K(x_n) = 2n + \mathcal{O}(\log n)$, $K(y_n) = 2n + \mathcal{O}(\log n)$, $I(x_n : y_n) = n + \mathcal{O}(\log n)$.

Theorem 3 *If z_n satisfies $K(z_n|x_n) = \mathcal{O}(\log n)$ and $K(z_n|y_n) = \mathcal{O}(\log n)$, then $K(z_n) = \mathcal{O}(\log n)$, that is* $\inf(\boldsymbol{x}, \boldsymbol{y}) = \Lambda$.

Proof. In the following we will omit the subscript in the expressions x_n, y_n, etc.

Lemma 2 *Let a bipartite graph $G = (V', V'', E)$ have no quadrangles, i.e. there are no $a, b \in V'$, $e, f \in V''$ such that $(a, e) \in E$, $(b, e) \in E$, $(a, f) \in E$, $(b, f) \in E$. Let $|V'| = |V''| = m$. Then $|E| \leq 1.5m^{1.5}$.*

Proof. Let $d = |E|/m$. Let V consist of $\sqrt{m}$ nodes from V' having maximum degree. The degree of any node in V is at least d. Any two different nodes in V have at most one common adjacent node in V'', as the graph has no quadrangles. Therefore, the set of nodes adjacent to the nodes in V is at least as big as the sum of degrees of nodes in V minus the number of pairs of nodes from V. Hence $m = |V''| \geq d\sqrt{m} - \binom{\sqrt{m}}{2} \geq d\sqrt{m} - m/2$. Thus, $d \leq 1.5\sqrt{m}$, $|E| = dm \leq 1.5m^{1.5}$. □

Assume that $K(z|x) = \mathcal{O}(\log n)$, $K(z|y) = \mathcal{O}(\log n)$, $K(z) = m$. Then

$$\begin{aligned} & K(x|z) = K(x, z) - K(z) + \mathcal{O}(\log n) = K(x) + K(z|x) - K(z) + \mathcal{O}(\log n) \\ = \; & K(x) - K(z) + \mathcal{O}(\log n) = 2n - m + \mathcal{O}(\log n). \end{aligned}$$

Analogously $K(y|z) = 2n - m + \mathcal{O}(\log n)$. Let $k = \max\{K(x|z), K(y|z)\}$.

[3] Actually, this is one of corollaries of Gaćs and Körner's theorem. The theorem itself is deeper

Let $V' = V''$ be the set of all lines l such that $K(l|z) \leq k$. Let E consist of all pairs of orthogonal lines in $V' \times V''$.

Then the bipartite graph (V', V'', E) has no quadrangles. By the Lemma 2, $|E| \leq 1.5 \cdot |V'|^{1.5} \leq 1.5 \cdot 2^{1.5(k+1)}$. As $(x,y) \in E$ we get $K(\langle x,y\rangle) \leq \log|E| + K(z) + \mathcal{O}(\log n) = 1.5k + m + \mathcal{O}(\log n) \leq 3n - 1.5m + m + \mathcal{O}(\log n)$. On the other hand, $K(\langle x,y\rangle) \geq 3n - \mathcal{O}(\log n)$. Hence $m = \mathcal{O}(\log n)$. □

4.3 Intersecting lines in the affine 3-dimensional space over $\mathbb{F}_q$

In this section the word "line" will mean an affine 1-dimensional subspace in the affine 3-dimensional space over $\mathbb{F}_q$. Let (x,y) be a random element in the set of all pairs of intersecting lines. We have $K(x) = 4n + \mathcal{O}(\log n)$, $K(y) = 4n + \mathcal{O}(\log n)$, $I(x:y) = n + \mathcal{O}(\log n)$.

Theorem 4 $\inf(\boldsymbol{x}, \boldsymbol{y}) = \Lambda$.

Proof. Assume that $K(z|x) = \mathcal{O}(\log n)$, $K(z|y) = \mathcal{O}(\log n)$. Let us prove that $K(z) = \mathcal{O}(\log n)$. Let $m = K(z)$. Then $K(x|z), K(y|z) = 4n - m + \mathcal{O}(\log n)$. Let B be the set of all lines $\hat{y}$ intersecting with x such that $K(\hat{y}|z) \leq K(y|z)$

Given z, x, we can enumerate all the lines in B and identify y by its number in that enumeration. Therefore $K(y|x) \leq K(z|x) + \log|B| + \mathcal{O}(\log n) = \log|B| + \mathcal{O}(\log n)$. On the other hand $K(y|x) = 3n + \mathcal{O}(\log n)$. Therefore, $\log|B| \geq 3n - \mathcal{O}(\log n)$ and $|B| \geq 2^{3n-\mathcal{O}(\log n)}$. Let us exclude from B all the lines having complexity less $3n - 2C\log n$ conditional to x, where C is the constant such that $|B| \geq 2^{3n - C\log n}$. Any line $\hat{y}$ in the resulting set B' intersects with x and has complexity $3n + \mathcal{O}(\log n)$ conditional to x. And $|B'| \geq 2^{3n-2C\log n}$.

Let $\hat{y}$ be any line in B'. We have $K(\langle x,\hat{y}\rangle) = K(x) + K(\hat{y}|x) + \mathcal{O}(\log n) = 7n + \mathcal{O}(\log n)$, $K(x|\hat{y}) = K(\langle x,\hat{y}\rangle) - K(\hat{y}) + \mathcal{O}(\log n) \geq 7n - 4n + \mathcal{O}(\log n) = 3n + \mathcal{O}(\log n)$. Since x and $\hat{y}$ intersect we have $K(x|\hat{y}) \leq 3n + \mathcal{O}(\log n)$, thus $K(x|\hat{y}) = 3n + \mathcal{O}(\log n)$. Finally, $K(\hat{y}) = K(\langle x,\hat{y}\rangle) - K(x|\hat{y}) + \mathcal{O}(\log n) = 7n - 3n + \mathcal{O}(\log n) = 4n + \mathcal{O}(\log n)$. And we have $K(z|\hat{y}) = K(z) + K(\hat{y}|z) - K(\hat{y}) + \mathcal{O}(\log n) \leq m + 4n - m - 4n + \mathcal{O}(\log n) = \mathcal{O}(\log n)$.

Let $A = A_{\hat{y}}$ be the set of all lines $\hat{x}$ intersecting with $\hat{y}$ such that $K(\hat{x}|z) \leq K(x|z)$. As $x \in A$ we have $K(x|\hat{y}) \leq \log|A| + \mathcal{O}(\log n)$. On the other hand $K(x|\hat{y}) = 3n + \mathcal{O}(\log n)$, so we get $|A| \geq 2^{3n-\mathcal{O}(\log n)}$.

Thus, for any of $2^{3n-\mathcal{O}(\log n)}$ lines $\hat{y}$ in B we have $|A_{\hat{y}}| \geq 2^{3n-\mathcal{O}(\log n)}$. Any line $\hat{x} \neq x$ belongs to about 2^{2n} different $A_{\hat{y}}$, as if $\hat{x} \in A_{\hat{y}}$, the line $\hat{y}$ intersects with both lines $\hat{x}$ and x and there are at most 2^{2n} lines intersecting with both $\hat{x}$ and x. Hence, the set $\bigcup_{\hat{y}\in B} A_{\hat{y}}$ has at least $\frac{2^{3n+\mathcal{O}(\log n)} \times 2^{3n+\mathcal{O}(\log n)}}{2^{2n}} = 2^{4n+\mathcal{O}(\log n)}$ elements. On the other hand, any line $\bigcup_{\hat{y}\in B} A_{\hat{y}}$ has complexity at most $4n - m + \mathcal{O}(\log n)$ conditional to z. Therefore, $m = \mathcal{O}(\log n)$. □

An interesting point in this example is that there exists a string of length $3n + \mathcal{O}(\log n)$ having $2n + \mathcal{O}(\log n)$ bits of mutual information both with x and y. (An evident upper bound for the complexity of such a z is $4n$.) Indeed, let z be the common point of x and y. Then $K(x|z) = 2n + \mathcal{O}(\log n)$, $K(y|z) = 2n + \mathcal{O}(\log n)$, therefore $I(x:z) \geq 2n + \mathcal{O}(\log n)$, $I(y:z) \geq 2n + \mathcal{O}(\log n)$, and $K(z) = 3n + \mathcal{O}(\log n)$.

The same holds for the plain containing both x and y.

4.4 Orthogonal lines in $\mathbb{F}_q^4$

In this section a line is a 1-dimensional subspace of the linear space $\mathbb{F}_q^4$. Let $x = x_n$ and $y = y_n$ be random independent orthogonal lines. Then $K(x) = 3n+\mathcal{O}(\log n)$, $K(y) = 3n+\mathcal{O}(\log n)$, $I(x:y) = n + \mathcal{O}(\log n)$.

Theorem 5 $\inf(x, y) = \Lambda$.

Proof. Let $K(z|x) = \mathcal{O}(\log n)$, $K(z|y) = \mathcal{O}(\log n)$. Denote $K(z)$ by m. Then $K(x|z) = 3n - m + \mathcal{O}(\log n)$, $K(y|z) = 3n - m + \mathcal{O}(\log n)$.

Let X [Y] be the set of all lines $\hat{x}$ [$\hat{y}$] such that $K(\hat{x}|z) \le K(x|z)$ [$K(\hat{y}|z) \le K(y|z)$]. It follows that

$$|X|, |Y| \le 2^{3n-m+\mathcal{O}(\log n)}.$$

On the other hand, since $x \in X$ and $K(x|z) = 3n-m+\mathcal{O}(\log n)$, we get $|X| \ge 2^{3n-m-\mathcal{O}(\log n)}$. Analogously, $|Y| \ge 2^{3n-m-\mathcal{O}(\log n)}$.

Lemma 3 *The number N of orthogonal pairs of lines in $X \times Y$ is at most* $2^{5n-1.5m+\mathcal{O}(\log n)}$

Proof. For any $\hat{x} \in X$ let $Y_{\hat{x}}$ be the set of lines in Y orthogonal to $\hat{x}$. Let $\tilde{X}$ be the set of $2^{n-\frac{m}{2}}$ lines $\hat{x}$ in X having maximum number of elements in $Y_{\hat{x}}$. Then the sum of cardinalities of $|Y_{\hat{x}}|$ for $\hat{x} \in \tilde{X}$ is at least $\frac{2^{n-\frac{m}{2}}N}{|X|}$. Any different $Y_{\hat{z}'}$ and $Y_{\hat{z}''}$ have at most $q+1 \approx 2^n$ common lines (as $Y_{\hat{z}'} \cap Y_{\hat{z}''}$ is a plain, and any plain has $q+1$ lines). Therefore the cardinality of the set $\bigcup_{\hat{x}\in\tilde{X}} Y_{\hat{x}}$ is at least

$$\frac{2^{n-\frac{m}{2}}N}{|X|} - \binom{|\tilde{X}|}{2} 2^n \ge \frac{2^{n-\frac{m}{2}}N}{|X|} - 2^{2n-m}2^n.$$

Thus,

$$2^{3n-m+\mathcal{O}(\log n)} \ge |Y| \ge \frac{2^{n-\frac{m}{2}}N}{|X|} - 2^{2n-m}2^n.$$

Hence $N \le 2^{3n-m+\mathcal{O}(\log n)}2^{-n+0.5m}|X| = 2^{5n-1.5m+\mathcal{O}(\log n)}$. □

Since the pair $\langle x, y\rangle$ is in $X \times Y$ and x is orthogonal to y, we get

$$K(\langle x, y\rangle) \le \log N + K(z) + \mathcal{O}(\log n) = 5n - 0.5m + \mathcal{O}(\log n).$$

As $K(\langle x, y\rangle) = 5n + \mathcal{O}(\log n)$, we get $m = \mathcal{O}(\log n)$. □

Though $\inf(\boldsymbol{x}, \boldsymbol{y}) = \Lambda$, there is a short string extracting much mutual information from x and y. Namely, there is a z such that $K(z) = 3n + \mathcal{O}(\log n)$, $K(x|z) = n + \mathcal{O}(\log n)$ and $K(y|z) = n + \mathcal{O}(\log n)$, that is $I(x : z) = 2n + \mathcal{O}(\log n)$ and $I(y : z) = 2n + \mathcal{O}(\log n)$. (An evident upper bound for the complexity of such a z is $4n$.)

The way to construct such a z was pointed by Finkelberg and Bezrukawnikow. Let W be the plain containing the vectors $(1, 0, 0, 0)$ and $(0, 1, 0, 0)$. Let w be any line in W

orthogonal to y (obviously it exists). Take as z the plain having the lines x and w (if it happens that $x = w$, take as z any plain containing x). Let us note that z has at least 1-dimensional intersection with W and the number of plains with this property is about q^3, therefore $K(z) \le \log q^3 + \mathcal{O}(\log n) = 3n + \mathcal{O}(\log n)$. The number of lines in z is at most $q+1$, thus $K(x|z) = \log q + \mathcal{O}(\log n) = n + \mathcal{O}(\log n)$. The line y is orthogonal to both x, w, therefore this line is orthogonal to z. The number of lines orthogonal to z is $q+1$, therefore $K(y|z) = \log q + \mathcal{O}(\log n) = n + \mathcal{O}(\log n)$.

The last two examples show that the greatest common lower bound is a rather rough tool to measure how much common information have x and y. In the next section we introduce a more precise tool to do this.

5 A finer approach

In this section we consider only the sequences $\boldsymbol{x}, \boldsymbol{y}$ such that $K(x_n)$, $K(y_n)$, $K(\langle x_n, y_n \rangle)$ have linear growth $(cn + \mathcal{O}(\log n))$.

Definition 2 *A triple of reals (u, v, w) is called $\boldsymbol{x}, \boldsymbol{y}$-admissible, if there exists a sequence $\boldsymbol{z}$ such that*

$$K(x_n|z_n) \le vn + \mathcal{O}(\log n),\ \ K(y_n|z_n) \le wn + \mathcal{O}(\log n),\ \ K(z_n) \le un + \mathcal{O}(\log n) \qquad (3)$$

(If we added the equations $un + vn = K(x_n) + \mathcal{O}(\log n)$, $un + wn = K(y_n) + \mathcal{O}(\log n)$, the inequalities (3) would mean that $\boldsymbol{z}$ is the common lower bound of $\boldsymbol{x}$ and $\boldsymbol{y}$.)

Let us denote the set of all $\boldsymbol{x}, \boldsymbol{y}$-admissible triples by $M_{\boldsymbol{x},\boldsymbol{y}}$. The larger is $M_{\boldsymbol{x},\boldsymbol{y}}$ the more common information have $\boldsymbol{x}$ and $\boldsymbol{y}$. Let, for instance, x_n be a random string of length n and $y_n = x_n$. In this case

$$M_{\boldsymbol{x},\boldsymbol{y}} = \{(u, v, w) \mid u + v \ge 1,\ u + w \ge 1\}.$$

If x_n, y_n are random independent strings of length n, then $M_{\boldsymbol{x},\boldsymbol{y}}$ is much smaller:

$$M_{\boldsymbol{x},\boldsymbol{y}} = \{(u, v, w) \mid u + v \ge 1,\ u + w \ge 1,\ u + v + w \ge 2\}.$$

For simplicity we restrict ourselves with $\boldsymbol{x}$, $\boldsymbol{y}$ such that

$$K(x_n) = 2n + \mathcal{O}(\log n),\ K(y_n) = 2n + \mathcal{O}(\log n),\ K(\langle x_n, y_n \rangle) = 3n + \mathcal{O}(\log n). \qquad (4)$$

Using the inequalities $K(\langle x_n, y_n \rangle) \le K(z_n) + K(x_n|z_n) + K(y_n|z_n) + \mathcal{O}(\log n)$ and $K(x_n) \le K(z_n) + K(x_n|z_n) + \mathcal{O}(\log n)$ it is easy to show that for all $\boldsymbol{x}, \boldsymbol{y}$-admissible triples it holds

$$u + v + w \ge 3,\ u + v \ge 2,\ u + w \ge 2. \qquad (5)$$

Thus, for every $\boldsymbol{x}, \boldsymbol{y}$ the set $M_{\boldsymbol{x},\boldsymbol{y}}$ is included in the set $M_{\max}$, defined by the inequalities (5).

Example 2 *Let $x_n = \langle p, q\rangle$, $y_n = \langle p, r\rangle$, where p, q, r are random independent strings of length n. It is easy to show that that the set of $\boldsymbol{x}, \boldsymbol{y}$-admissible triples is equal to $M_{\max}$. This fact agrees with our intuition that $\boldsymbol{x}$ and $\boldsymbol{y}$ have as much common information as possible (under restriction (4)).*

Let us define $M_{\min}$ to be the set of all triples u, v, w satisfying inequalities (5) and *at least one* of the following inequalities:

$$u + v + w \geq 4,\ \ u + v \geq 3,\ \ u + w \geq 3. \tag{6}$$

Theorem 6 *For all $\boldsymbol{x}, \boldsymbol{y}$ satisfying (4) it holds $M_{\min} \subseteq M_{\boldsymbol{x},\boldsymbol{y}} \subseteq M_{\max}$.*

Proof. Let $\boldsymbol{x}, \boldsymbol{y}$ satisfy (4). We have already shown that $M_{\boldsymbol{x},\boldsymbol{y}} \subseteq M_{\max}$.

Let us prove that $M_{\min} \subseteq M_{\boldsymbol{x},\boldsymbol{y}}$. Let (u, v, w) be in $M_{\min}$. Then the triple (u, v, w) satisfies the inequalities (5) and at least one of the inequalities (6). So consider three cases.

1) $u + v + w \geq 4$. If $v, w \leq 2$ let z be the concatenation of the first $(2-v)n$ bits of x and the first $(2 - w)n$ bits of y. Since $u + v + w \geq 4$, we have $|z| = (2 - v)n + (2 - w)n \leq un$. To obtain x given z we need the remaining vn bits of x and the numbers n, vn, wn, so $K(x|z) \leq vn + \mathcal{O}(\log n)$. Analogously, $K(y|z) \leq wn + \mathcal{O}(\log n)$.

Otherwise, if say $v > 2$, let z consist of the first $\min\{2, u\}n$ bits of y. Then $K(y|z) \leq (2 - \min\{2, u\})n + \mathcal{O}(\log n) \leq wn + \mathcal{O}(\log n)$, as the triple (u, v, w) satisfies (5). And $K(x|z) \leq K(x) \leq 2n + \mathcal{O}(\log n) \leq vn + \mathcal{O}(\log n)$.

2) $u+v \geq 3$. If $u \leq 2$ let z consist of the first un bits of y. To find x given z is suffices to know the remaining $(2-u)n$ bits of y and the minimum program to compute x given y (having n bits). So the total number of bits needed to find x given u is $(2 - u)n + n + \mathcal{O}(\log n) \leq vn + \mathcal{O}(\log n)$. And $K(y|z) \leq (2 - u)n + \mathcal{O}(\log n) \leq wn + \mathcal{O}(\log n)$.

Otherwise (if $u > 2$) let z be the concatenation of y and the first $\min\{u - 2, 1\}n$ bits of minimum program p to compute x given y. To obtain x given z it suffices to have the remaining $n - (u - 2)n \leq vn$ bits of p.

3) $u + w \geq 3$. Similar to 2). □

It turns out that there exists $\boldsymbol{x}, \boldsymbol{y}$ such that $M_{\boldsymbol{x},\boldsymbol{y}} = M_{min}$. This can by proved using the method of [7].

Theorem 7 *There are sequences $\boldsymbol{x}$, $\boldsymbol{y}$ such that $M_{\boldsymbol{x},\boldsymbol{y}} = M_{\min}$.*

Proof. The proof easily deduces from the following lemma.

Lemma 4 *There are $\boldsymbol{x}, \boldsymbol{y}$ satisfying (5) such that for any n there is no $\boldsymbol{z}$ satisfying the inequalities*

$$\begin{aligned} K(z_n) + K(x_n|z_n) + K(y_n|z_n) &\leq 4n && (7)\\ K(z_n) + K(x_n|z_n) &\leq 3n && (8)\\ K(z_n) + K(y_n|z_n) &\leq 3n. && (9) \end{aligned}$$

Proof. Let us fix natural n. As usually we will omit the subscript n in x_n, y_n, etc.

Let U be the set of all strings of length $2n + C\log n$, where constant C will be chosen later. Let

$$\begin{aligned} U_1 &= \{u \in U \mid K(u) < 2n\} \\ V &= \{(x, y) \mid x, y \in U,\ K(\langle x, y\rangle) < 3n\} \\ V_1 &= \{(x, y) \mid x, y \in U,\ \text{there is } c \text{ satisfying the inequalities (7), (8), and (9)}\}. \end{aligned}$$

We will show that the set $(U \times U) \setminus [(U_1 \times U_1) \cup V \cup V_1]$ is non-empty. Any pair (x, y) in this set will satisfy the following:
1) $K(x), K(y) = 2n + \mathcal{O}(\log n)$ (as both x and y are in $U \setminus U_1$),
2) $K(\langle x, y\rangle) \geq 3n$ (as $\langle x, y\rangle \notin V$), and
3) there is no z satisfying the inequalities (7), (8), and (9) (as $\langle x, y\rangle \notin V_1$).

Thus, to prove the lemma it suffices to show that there is (x, y) in $(U \times U) \setminus [(U_1 \times U_1) \cup V \cup V_1]$ of complexity at most $3n + \mathcal{O}(\log n)$.

The non-emptyness of $(U \times U) \setminus [(U_1 \times U_1) \cup V \cup V_1]$ is proved by counting arguments. We have $|U| = 2^{4n}n^C$, $|U_1| < 2^{2n}$, $|V| < 2^{3n}$. To obtain an upper bound for $|V_1|$ let us count the number of pairs (x, y) for which there is z satisfying the inequality (7). For any k, l, m there are at most $2^k 2^l 2^m$ pairs x, y such that there is z with $K(z) = k$, $K(x|z) = l$, $K(y|z) = m$. And the number of triples k, l, m satisfying the inequality $k + l + m \leq 4n$ is at most $(4n + 1)^3$. Therefore, $|V_1| \leq (4n + 1)^3 2^{4n}$. It follows that if C is big enough, then $|U| = 2^{4n}n^C > 2^{2(2n)} + 2^{3n} + (4n + 1)^3 2^{4n} \geq |U_1 \times U_1| + |V| + |V_1|$, and therefore the set $(U \times U) \setminus [(U_1 \times U_1) \cup V \cup V_2]$ is non-empty.

Let (x, y) be the lexicographically first pair in $(U \times U) \setminus [(U_1 \times U_1) \cup V \cup V_2]$.

Lemma 5 $K(\langle x, y\rangle) \leq 3n + \mathcal{O}(\log n)$.

Proof. To identify x, y it suffices to know n and the sets U_1, V and V_1.

Let ϕ_0 be the universal conditional description method. For any $k + l \leq 3n$ let $W_{k,l}$ be the set of all (p, q) such that $|p| = k$, $|q| = l$ and both $\phi_0(p, \epsilon)$ and $G(\phi_0(p, \epsilon), q)$ are defined. To identify V_1 it suffices to know n and the sets $W_{k,l}$ for all k, l such that $k + l \leq 3n$.

Therefore, x, y can be retrieved from n and the sets U_1, V and $W_{k,l}$, $k + l \leq 3n$.

The elements of all the sets U_1, V and $W_{k,l}$ can be enumerated given n, therefore to get the lists of all these sets it suffices to know n and the number $m = |U_1| + |V| + \sum_{k+l\leq 3n} |W_{k,l}|$ (given n we enumerate elements in all these sets until m elements are enumerated). We have

$$|U_1| \leq 2^{2n},\ |V| \leq 2^{3n},\ |W_{k,l}| \leq 2^k 2^l \leq 2^{3n}.$$

Therefore

$$|U_1| + |V| + \sum_{k+l\leq 3n} |W_{k,l}| \leq (3n + 3)2^{3n},$$

and

$$K(\langle x, y\rangle) \leq \log(|U_1| + |V| + \sum_{k+l\leq 3n} |W_{k,l}|) + 2\log n + C \leq 3n + \mathcal{O}(\log n). \ \square$$

This finishes the proof of Lemma 5. □

We claim that $M_{\boldsymbol{x},\boldsymbol{y}} = M_{\min}$ for any sequence satisfying Lemma 4. Assume for the contrary that the set $M_{\boldsymbol{x},\boldsymbol{y}} \setminus M_{\min}$ is not empty, that is there is a triple (u, v, w) satisfying the inequalities

$$u + v + w < 4, \ u + v < 3, \ u + w < 3,$$

for which there exists a sequence $\boldsymbol{z}$ satisfying (3). Then for n large enough we get

$$\begin{aligned} K(z_n) + K(x_n|z_n) + K(y_n|z_n) &\leq un + vn + wn + \mathcal{O}(\log n) < 4n \\ K(z_n) + K(x_n|z_n) &\leq un + vn + \mathcal{O}(\log n) < 3n \\ K(z_n) + K(y_n|z_n) &\leq un + wn + \mathcal{O}(\log n) < 3n. \end{aligned}$$

The contradiction shows that $M_{\boldsymbol{x},\boldsymbol{y}} = M_{\min}$. □

The proof of Theorem 7 is non-constructive, it gives no "example" of the pair $(\boldsymbol{x}, \boldsymbol{y})$ with $M_{\min} = M_{\boldsymbol{x},\boldsymbol{y}}$. An example would be a set A_n of low complexity ($\mathcal{O}(\log n)$) such that any random pair (x_n, y_n) in this set satisfies Theorem 7. We do not know whether such a proof exists. In the Section 4, we presented several examples (in this sense) of sequences whose common information is less than mutual information. However we do not know whether the set $M_{\boldsymbol{x},\boldsymbol{y}}$ is minimal for these $\boldsymbol{x}, \boldsymbol{y}$. To be more precise, we should note that the set $M_{\boldsymbol{x},\boldsymbol{y}}$ for $\boldsymbol{x}, \boldsymbol{y}$ defined in Sections 4.2, 4.3, 4.4 might depend on the choice of q. We will present later some arguments in favor of this possibility.

Now, we establish an upper bound for $M_{\boldsymbol{xy}}$ for $(\boldsymbol{x}, \boldsymbol{y})$ being a random pair of orthogonal lines in $\mathbb{F}_q^3$ (see Section 4.2), which holds for any choice of q. Denote by $\tilde{M}$ be the set of all $\boldsymbol{x}, \boldsymbol{y}$-admissible triples. Let

$$\hat{M} = \{\langle u, v, w\rangle \mid u + v/2 + \max\{w, v/2\} \geq 3, \ u + w/2 + \max\{v, w/2\} \geq 3\} \cap M_{\max}.$$

Note that the set $\hat{M}$ is a proper subset of $M_{\max}$ and a proper superset of $M_{\min}$. For instance, the triple $(1.5, 1, 1)$ is in $\hat{M} \setminus M_{\min}$ and the triple $(1, 1, 1)$ is in $M_{\max} \setminus \hat{M}$.

Theorem 8 $\tilde{M} \subseteq \hat{M}$.

Proof. Let us fix n. As usual, we omit subscript n.

Let $\langle u, v, w\rangle \in \tilde{M}$, that is there exists $\boldsymbol{z}$ such that the inequalities (3) hold.

We need the following extension of Lemma 2, which is proved by the same arguments.

Lemma 6 *Let a bipartite graph $G = (V', V'', E)$ have no quadrangles. Then for any integer a not greater than $|V'|$ it holds $\frac{a|E|}{|V'|} - a^2 \leq |V''|$.*

(In Lemma 2 we had $|V'| = |V''| = m$, $a = \sqrt{m}$.)

Let V' $[V'']$ be the set of all lines having complexity at most $K(x|z)$ $[K(y|z)]$ conditional to z. Put an edge between $\hat{x} \in V'$ and $\hat{y} \in V''$ if $\hat{x}$ is orthogonal to $\hat{y}$. As (x, y) is in E and the elements in E can be enumerated given q and z, we get

$$3n \leq K(\langle x, y\rangle) \leq \log |E| + K(z) + \mathcal{O}(\log n).$$

By Lemma 6 we get
$$\frac{a2^{3n-K(z)-\mathcal{O}(\log n)}}{|V'|}-a^2\le |V''|,$$
or
$$\frac{a2^{3n}}{|V'|}\le 2^{K(z)+\mathcal{O}(\log n)}(a^2+|V''|)$$
for any $a\le |V'|$.

If $\sqrt{|V''|}\le |V'|$ we let $a=\sqrt{|V''|}$. We get then
$$\frac{\sqrt{|V''|}2^{3n}}{|V'|}\le 2^{K(z)+\mathcal{O}(\log n)}\cdot 2|V''|,$$
or
$$2^{3n}\le 2^{K(z)+\mathcal{O}(\log n)}\cdot 2\sqrt{|V''|}|V'|.$$
If $\sqrt{|V''|}\ge |V'|$ we let $a=|V'|$. We get then
$$2^{3n}\le 2^{K(z)+\mathcal{O}(\log n)}(|V'|^2+2|V''|),$$
thus
$$2^{3n}\le 2^{K(z)+\mathcal{O}(\log n)}\cdot 3|V''|.$$
Thus, anyway we have
$$2^{3n}\le 2^{K(z)+\mathcal{O}(\log n)}\cdot 3\sqrt{|V''|}\max\{\sqrt{|V''|},|V'|\}.$$
The number of elements in V' and V'' is at most $2^{K(x|z)+1}$ and $2^{K(y|z)+1}$, respectively, and $K(z)\le un+\mathcal{O}(\log n)$, $K(x|z)\le vn+\mathcal{O}(\log n)$, $K(y|z)\le wn+\mathcal{O}(\log n)$. Therefore, $3n\le un+0.5wn+\max\{0.5w,v\}n+\mathcal{O}(\log n)$, thus $3\le u+0.5w+\max\{0.5w,v\}$.

In the similar way we can prove that $3\le u+0.5v+\max\{0.5v,w\}$. $\square$

Let us explain now why we think the set $\tilde{M}$ might depend on the choice of q. Suppose that q is a square, $q=p^2$ (for all n). Then we claim that the set $M_{\boldsymbol{x},\boldsymbol{y}}$ has the point $(1.5,1,1)$, which is on the border of $\hat{M}$.

Let $\alpha\in\mathbb{F}_q$ be a primitive element of $\mathbb{F}_q$ over $\mathbb{F}_p$. Thus any element in F_q can be represented in the form $t+s\alpha$ for some $t,s\in\mathbb{F}_p$. We can choose α in such a way that, moreover, any element in F_q can be represented in the form $t+s\alpha^2$ for some $t,s\in\mathbb{F}_p$. Why? The multiplicative group of the field $\mathbb{F}_q$ is cyclic (see [4, page 184]), therefore the square of any its generator does not belong to $\mathbb{F}_p$. Let us take as α any such generator. Then $\alpha^2=e+f\alpha$, where $e,f\in\mathbb{F}_p$ and $f\ne 0$. Thus α is a linear combination of $1,\alpha^2$ with coefficients from $\mathbb{F}_p$, and we are done.

Let us find z of complexity $1.5n+\mathcal{O}(\log n)$ such that $K(x|z)=n+\mathcal{O}(\log n)$, $K(y|z)=n+\mathcal{O}(\log n)$. Let (a,b,c) be the leading vector of x (defined up to a multiplicative constant). We may assume that $c\ne 0$, since the number of lines for which $c=0$ is equal to $q+1$, therefore the complexity of any such line is at most $\log(q+1)+\mathcal{O}(\log n)\le n+\mathcal{O}(\log n)$.

So let $c = 1$. By the same reason we may assume that the leading vector of the line y is $(a', 1, c')$. As y is orthogonal to x, we get $c' = -(aa' + b)$.

We have $a = z_1 + r\alpha$, $a' = z_2 + t\alpha$, where $z_1, z_2, r, t \in \mathbb{F}_p$. Find $z_3, s \in \mathbb{F}_p$ such that $b = -z_2 r\alpha + z_3 + s\alpha^2$. This is possible by our assumption on α. Let $z = \langle z_1, z_2, z_3\rangle$. Obviously, $K(z) \leq 3\log p + \mathcal{O}(\log n) = 1.5n + \mathcal{O}(\log n)$. Given z, r and s we can find x, therefore $K(x|z) \leq K(r) + K(s) + \mathcal{O}(\log n) \leq 2\log p + \mathcal{O}(\log n) \leq n + \mathcal{O}(\log n)$.

Let us prove that $K(y|z) \leq n + \mathcal{O}(\log n)$. It is easy to see that

$$c' = -(z_1 z_2 + z_1 t\alpha + rt\alpha^2 + z_3 + s\alpha^2) = -(z_1 z_2 + z_1 t\alpha + z_3 + (rt + s)\alpha^2).$$

Therefore, given z, t and $rt+s$ we can find y. Hence $K(y|z) \leq K(t) + K(rt+s) + \mathcal{O}(\log n) \leq 2\log p + \mathcal{O}(\log n) \leq n + \mathcal{O}(\log n)$.

So, if we let for instance, $q_n = 2^{2[n/2]}$ we result with $\boldsymbol{x}, \boldsymbol{y}$ for which the set $\tilde{M}$ has the point $(1.5, 1, 1)$. And we do not know whether this is the case for (say) $q = 2^{2[n/2]+1}$.

A natural related question is the following one.

Question 1. Does there exist a sequence $G_n = (V', V'', E)$ of bipartite graphs having no quadrangles for which $K(G_n) = \mathcal{O}(\log n)$, $|V'|, |V''| = 2^{2n+\mathcal{O}(\log n)}$, $|E| = 2^{3n+\mathcal{O}(\log n)}$ and $M_{\boldsymbol{xy}} = \hat{M}$, where (x_n, y_n) is any random edge in E.

6 Some properties of the structure $\langle R, \leq\rangle$.

1. The operations of taking infimum and supremum are not distributive, even when they are defined. Indeed, let y_n and z_n be random independent strings of length n and let $x_n = y_n \oplus z_n$. Then

$$\sup(\inf(\boldsymbol{x}, \boldsymbol{y}), \boldsymbol{z}) \neq \inf(\sup(\boldsymbol{x}, \boldsymbol{z}), \sup(\boldsymbol{y}, \boldsymbol{z})),$$

as

$$\sup(\inf(\boldsymbol{x}, \boldsymbol{y}), \boldsymbol{z}) = \sup(\Lambda, \boldsymbol{z}) = \boldsymbol{z},$$

and

$$\inf(\sup(\boldsymbol{x}, \boldsymbol{z}), \sup(\boldsymbol{y}, \boldsymbol{z})) = \inf(\langle \boldsymbol{y}, \boldsymbol{z}\rangle, \langle \boldsymbol{y}, \boldsymbol{z}\rangle) = \langle \boldsymbol{y}, \boldsymbol{z}\rangle.$$

Besides that,

$$\inf(\sup(\boldsymbol{x}, \boldsymbol{y}), \boldsymbol{z}) \neq \sup(\inf(\boldsymbol{x}, \boldsymbol{z}), \inf(\boldsymbol{y}, \boldsymbol{z})),$$

as

$$\inf(\sup(\boldsymbol{x}, \boldsymbol{y}), \boldsymbol{z}) = \inf(\langle \boldsymbol{y}, \boldsymbol{z}\rangle, \boldsymbol{z}) = \boldsymbol{z}$$

and

$$\sup(\inf(\boldsymbol{x}, \boldsymbol{z}), \inf(\boldsymbol{y}, \boldsymbol{z})) = \sup(\Lambda, \Lambda) = \Lambda.$$

2. Any two elements in R have their difference. (A sequence $\boldsymbol{z}$ is called a difference of $\boldsymbol{x}$ and $\boldsymbol{y}$, if $\sup(\boldsymbol{y}, \boldsymbol{z}) = \sup(\boldsymbol{y}, \boldsymbol{x})$ and $\inf(\boldsymbol{y}, \boldsymbol{z}) = \Lambda$.) Indeed, let z_n be a minimum program to compute x_n given y_n. It is easy to verify that $\sup(\boldsymbol{y}, \boldsymbol{z}) = \sup(\boldsymbol{y}, \boldsymbol{x})$ and $\inf(\boldsymbol{y}, \boldsymbol{z}) = \Lambda$.

The difference is not defined uniquely: let y_n and z_n be random independent strings of length n, both y_n and $y_n \oplus z_n$ are differences of y_n and z_n.

7 Conclusion

Some questions remain unresolved. One of them is mentioned in the previous section. Here are the others.

Question 2.[4] Let us modify the relation "$\boldsymbol{x}$ is easy given $\boldsymbol{y}$" as follows. Let us fix a function $f(n) = o(n)$ and let

$$\boldsymbol{x} \leq_f \boldsymbol{y} \rightleftharpoons K(x_n|y_n) = O(f(n)).$$

We do not know whether any two sequences in R have the greatest common lower bound with respect to $\leq_f$? We conjecture that the answer is negative for all $f(n) = o(n)$ (as it was shown for $f(n) = \log n$).

The same question for the pre-order defined by

$$\boldsymbol{x} \leq \boldsymbol{y} \rightleftharpoons K(x_n|y_n) = o(n).$$

Question 3. Find good lower and upper bound for $M_{\boldsymbol{x},\boldsymbol{y}}$ for pairs $(\boldsymbol{x}, \boldsymbol{y})$ defined in Sections 4.1, 4.2, 4.3 and 4.4.

Question 4. Denote by $\mathcal{R}$ the set R factored over the equivalence relation $\boldsymbol{x} \leq \boldsymbol{y} \wedge \boldsymbol{y} \leq \boldsymbol{x}$. Let $[\boldsymbol{x}]$ denote the factor class containing $\boldsymbol{x}$. Let us call a bijection $f : \mathcal{R} \to \mathcal{R}$ an *automorphism* if $K(f(x_n)|f(y_n)) = K(x_n|y_n) + \mathcal{O}(\log n)$ (here, $f(x_n), f(y_n)$ denote any sequences in $f([\boldsymbol{x}]), f([\boldsymbol{y}])$, respectively; it is easy to verify that this definition is correct). Does there exist, for any $\boldsymbol{x}, \boldsymbol{y}$ such that $K(x_n) = K(y_n) + \mathcal{O}(\log n)$, an automorphism f such that $f([\boldsymbol{x}]) = [\boldsymbol{y}]$?

References

[1] G.J. Chaitin. "On the length of programs for computing finite binary sequences." *J. Assoc. Comput. Mach.*, vol. 16 (1969), pp. 145–159.

[2] P. Gács, J. Körner. "Common information is far less than mutual information." *Problems of Control and Information Theory,* vol. 2 (1973), pp. 149–162.

[3] A.N. Kolmogorov. "Three approaches to the quantitative definition of information". *Problems Inform. Transmission*, vol. 1 (1965), No. 1, pp. 1–7.

[4] S. Lang. Algebra. Addison-Wesley, 1965.

[5] M. Li, P. Vitányi. An introduction to Kolmogorov complexity and its applications. Springer, 1997.

[6] An.A. Muchnik, "On the extraction of common information of two strings." Abstracts of talks at the First World Congress of Bernoulli Society, Moscow, Nauka, 1986, p. 453.

[4] Recently answered by the second author

[7] An.A. Muchnik. "Common information". Manuscript, 1996.

[8] R.J. Solomonoff. "A formal theory of inductive inference, part 1 and part 2". *Inform. Control*, vol. 7 (1964), pp. 1–22, 224–254.

[9] V. A. Uspensky, A. Shen. "Relations between varieties of Kolmogorov complexities." *Mathematical Systems Theory*, vol. 29 (1996), no. 3, pp. 271–292.

[10] A.K. Zvonkin, L.A. Levin, "The complexity of finite objects and the development of the concepts of information and randomness by means of the theory of algorithms." *Russian Math. Surveys*, vol. 25 (1970), No. 6. pp. 83–124.