ELSEVIER

Theoretical Computer Science 207 (1998) 319–328

Theoretical Computer Science

On common information

An.A. Muchnik *

Institute of New Technologies, Nizhnaya Radishevskaya 10, 109004 Moscow, Russia

Abstract

We present a complexity-theoretic proof of the result of P.Gács and J.Körner on the existence of a pair of words whose common information can not be materialized. Our method is easier than Gács and Körner's method and gives a possibility to get some generalization of the result. Besides, we show that there are many enough pairs of words with this property.

Keywords: Kolmogorov entropy; Algorithmic information theory; Common information of two words

1. Introduction

The results of the present paper were announced in [3].

In paper [2] Kolmogorov gave the definition of the entropy $K(x)$ of finite object x and the definition of the quantity of information $I(x : y)$ about y contained in x. In [4] it was proved that the quantity of information is commutative to within an additive term of $O(\log K(x) + \log K(y))$ (all logarithms in our paper have base two). In that paper it was also proved that

$$|I(x : y) - (K(x) + K(y) - K(x, y))| = O(\log K(x) + \log K(y)).$$

Because of commutativity of $I(x : y)$ we will call $I(x : y)$, $I(y : x)$ and $K(x) + K(y) - K(x, y)$ the quantity of common information of x, y.

My father, A.A. Muchnik, raised the following question: is there a word z such that $K(z) = I(z : x) = I(z : y) = I(x : y)$? That is, can we materialize the common information? As $I(x : y)$ is commutative only to within an additive term of $O(\log K(x) + \log K(y))$ the exact formulation of this question is the following.

The problem of A.A. Muchnik. Is the following assertion true?

* Fax: 7095 915-69-63; e-mail: amuchnik@int.glasnet.ru.

0304-3975/98/$19.00
PII S0304-3975(98)00070-X

For some constant c for all binary words x, y there is a binary word z such that

$$\begin{aligned} &K(z)\leqslant I(x:y)+c(\log K(x)+\log K(y))+c,\\ &I(z:x)\geqslant I(x:y)-c(\log K(x)+\log K(y))-c,\\ &I(z:y)\geqslant I(x:y)-c(\log K(x)+\log K(y))-c. \end{aligned} \tag{1}$$

In paper [1] Gács and Körner gave a negative solution to this problem. They used probabilistic methods. In the present paper we present a complexity-theoretic solution to the problem of A.A. Muchnik, which is easier than Gács and Körner's solution. We are interested also in the following generalization of the question. Let us fix the parameters $m,n,a,b,i\in\mathbb{N}$ such that $a\leqslant m$, $b\leqslant n$, $i\leqslant m$, $i\leqslant n$.

What is the minimal d such that for every words x, y which have, respectively, the entropies m, n and the quantity of common information i there is a word z such that $K(z)\leqslant d$, $I(z:x)\geqslant a$, $I(z:y)\geqslant b$? We can easily prove that

$$d\leqslant a+b+2(\log m+\log n)+\text{const}.$$

It turns out that this upper bound is tight (to within an additive logarithmic term) if $m\geqslant a+i$, $n\geqslant b+i$ (Theorem 2).

We are also interested in the following question: for how many x, y there is no z such that the assertion (1) is true? We give the following answer. Let us fix some value of c in (1). Then for all sufficiently large n for all x if $K(x)\geqslant n$ then there is y such that $K(y)$ is equal to $K(x)$ to within an additive logarithmic term and such that there is no z satisfying (1) (Theorem 3).

2. Basic definitions

Let us denote by Ξ the set of all binary words. Let us denote the length of a word x by $l(x)$. Let us call any partial computable function $f\colon \Xi\longrightarrow\Xi$ the specifying method. We call the word p a description of x (with respect to f) if $f(p)=x$.

The complexity $K_f(x)$ of word x with respect to the specifying method f is defined as

$$K_f(x)=\min\{l(p)\mid f(p)=x\}$$

and $K_f(x)=\infty$ if there is no p such that $f(p)=x$. We say that specifying method f is no worse than a specifying method g if there is some $c\in\mathbb{N}$ such that for all $x\in\Xi$ $K_f(x)\leqslant K_g(x)+c$. The well known theorem of Solomonoff–Kolmogorov states that there is an *optimal* specifying method, that is a method which is no worse than all other ones. Let us fix some optimal specifying method and denote it by f_0. Let us denote $K(x)=K_{f_0}(x)$. We call $K(x)$ the *entropy* of x. Obviously, for some constant c all x satisfy the inequality $K(x)\leqslant l(x)+c$. Therefore, $K(x)<\infty$ for all x.

In the above inequality, c denotes a constant in the following sense: c doesn't depend on x but c depends on the choice of f_0. In our paper there are also absolute constants. We will give absolute constants explicitly.

Let g: $\Xi \times \Xi \longrightarrow \Xi$ be partial computable function. The complexity K_g of x, with respect to g, conditional to y is defined by

$$K_g(x|y) = \min\{l(p) \mid g(p,y) = x\}.$$

Among all partial computable functions g: $\Xi \times \Xi \longrightarrow \Xi$ there is an optimal one, that is a function g such that for all h there is a constant $c \in \mathbb{N}$ for which $K_g(x|y) \leqslant K_h(x|y) + c$ for all $x, y \in \Xi$. Let us fix some optimal g and denote it by g_0. We define $K(x|y) = K_{g_0}(x|y)$. Let us call the difference

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

the quantity of information about y contained in x.

Let us define some convenient coding of pairs of binary words. Let x be a word, $x = b_1 b_2 \ldots b_n$, $b_i \in \{0,1\}$; denote by $\bar{x}$ the word $b_1 b_1 b_2 b_2 \ldots b_n b_n 01$ (for example, $\overline{010} = 00110001$). Obviously given $\bar{x}y$ we can find both x and y. Let us denote $K(\bar{x}y)$ by $K(x,y)$. The well known theorem ([4]) states that there is $c \in \mathbb{N}$ such that for every $x, y \in \Xi$

$$|K(x,y) - (K(x) + K(y|x))| \leqslant 5(\log K(x) + \log K(y)) + c.$$

This yields

$$|I(x : y) - (K(x) + K(y) - K(x,y))| \leqslant 5(\log K(x) + \log K(y)) + c,$$

$$|I(x : y) - I(y : x)| \leqslant 10(\log K(x) + \log K(y)) + c_1.$$

Let us denote for the sake of convenience $(K(x) + K(y) - K(x,y))$ by $\bar{I}(x : y)$. We will write $a =_c b$ instead of $|a - b| < c$.

3. The results

Let $a, b \in \mathbb{N}$; $x, y \in \Xi$. Let $C_{a,b}(x,y) = \min\{K(z) \mid I(z : x) \geqslant a, I(z : y) \geqslant b\}$. The following simple theorem yields upper and lower bounds for $C_{a,b}(x,y)$. By m, n, i we denote respectively $K(x), K(y), \bar{I}(x : y)$.

Theorem 1. *There is a constant $c \in \mathbb{N}$ such that if $m \geqslant a + c$, $n \geqslant b + c$ then*

$$\max\{a, b, a + b - i\} - 40(\log m + \log n) - c \leqslant C_{a,b}(x,y) \leqslant a + b + 2(\log m + \log n) + c.$$

Proof. We will denote in this proof as well as in others proofs by $c_1, c_2, \ldots$ constants depending only on the choice of f_0, g_0. The value of c will be clear from the proof.

Let us prove first the lower bound of $C_{a,b}(x,y)$. We have to prove three inequalities:

$$a - 40(\log K(x) + \log K(y)) - c \leqslant C_{a,b}(x,y),$$
$$b - 40(\log K(x) + \log K(y)) - c \leqslant C_{a,b}(x,y),$$
$$a + b - \bar{I}(x:y) - 40(\log K(x) + \log K(y)) - c \leqslant C_{a,b}(x,y).$$

The first two inequalities follow from the well known inequality $K(x) \leqslant K(x|z) + K(z) + 2\log K(x) + c_1$, i.e. $K(z) \geqslant I(z:x) - 2\log K(x) - c_1$.

Let us prove the third inequality. This inequality follows from the inequality

$$K(z) \geqslant I(z:x) + I(z:y) - \bar{I}(x:y) - 40(\log K(x) + \log K(y)) - c_2.$$

Let us prove this inequality. Note that we can take as z the pair (x,y). Therefore $K(z) \leqslant K(x) + K(y) + 2\log K(x) + c_1$. Since the function log is concave, this implies that $\log K(z) \leqslant 2(\log K(x) + \log K(y)) + c_1$. In the next proof we will omit additive logarithmic terms.

$$\begin{aligned} I(z:x) + I(z:y) - \bar{I}(x:y) &= K(x) + K(z) - K(x,z) + K(y) + K(z) \\ &\quad - K(y,z) - K(x) - K(y) + K(x,y) \\ &= 2K(z) + K(x,y) - K(x,z) - K(y,z). \end{aligned}$$

As $K(x,z) = K(x|z) + K(z)$ we get

$$\begin{aligned} K(z) + K(x,y) - K(x,z) - K(y,z) &= K(x,y) - K(x|z) - K(y,z) \\ &\leqslant K(x,y) - K(x|\bar{y}z) - K(\bar{y}z) \\ &= K(x,y) - K(x,\bar{y}z) \leqslant 0. \end{aligned}$$

Thus to within an additive logarithmic term

$$I(z:x) + I(z:y) - \bar{I}(x:y) \leqslant K(z).$$

Now let us prove the upper bound of $C_{a,b}$. Clearly, we can take as z the concatenation of a first bits of the shortest description of x, b first bits of the shortest description of y and $\overline{l(a)}$ ($\overline{l(a)}$ stands for $\bar{u}$, where u is binary representation of $l(a)$).

The theorem is proved.

It is easy to prove that the lower bound in Theorem 1 is tight (to within an additive logarithmic term). Indeed, let $m \geqslant a$, $n \geqslant b$, $m \geqslant i$, $n \geqslant i$. Let us define words x, y such that $K(x) = m$, $K(y) = n$, $\bar{I}(x:y) = i$, $C_{a,b}(x,y) = \max\{a, b, a + b - i\}$ (factors of logarithmic length in words and logarithmic terms in numbers are omitted in the following reasoning). Let $a \geqslant b$. Let us take random mutually independent words p, q, r with lengths respectively i, $m - i$, $n - i$, i.e. $K(p|\bar{q}r) = i$, $K(q|\bar{p}r) = m - i$, $K(r|\bar{p}q) = n - i$. Let $x = pq$, $y = pr$. We consider two cases.

(1) $a \geqslant a + b - i$, i.e. $i \geqslant b$.

Let z be the beginning of x of length a. Then $I(z:x) = a$. If $i \geqslant a$ then z is beginning of y consequently $I(z:y) = a \geqslant b$. If $a \geqslant i$ then z begins with p therefore $I(z:y) = i \geqslant b$.

(2) $a+b-i \geqslant a$, i.e. $b \geqslant i$.

Let z be equal to pq_1r_1, where q_1, r_1 are respectively the beginnings of q, r of lengths $a-i$, $b-i$. Then $K(z) = i+a-i+b-i = a+b-i$, $I(z:x) = i+a-i = a$, $I(z:y) = i+b-i = b$.

The following theorem shows that the upper bound in Theorem 1 is tight if $m \geqslant a+i$, $n \geqslant b+i$.

Theorem 2. *There is a constant $c \in \mathbb{N}$ such that the following holds. For every m, n, a, b, i such that $m \geqslant i$, $n \geqslant i$, $m \geqslant a$, $n \geqslant b$ there are x, y such that $K(x) =_c m$, $K(y) =_c n$, $|\bar{I}(x:y) - i| \leqslant 5(\log m + \log n) + c$, $C_{a,b}(x,y) \geqslant \max\{a, b, \min\{a+n-i, b+m-i, a+b\}\} - 5(\log m + \log n) - c$.*

Proof. Let m, n, a, b, i satisfy the inequalities $m \geqslant i$, $n \geqslant i$, $m \geqslant a$, $n \geqslant b$. Let us denote $k = \min\{a+b, a+n-i, b+m-i\}$.

It is easy to see that $C_{a,b}(x,y) \geqslant a - 5(\log m + \log n) - c$ and $C_{a,b}(x,y) \geqslant b - 5(\log m + \log n) - c$. So we have to prove that $C_{a,b}(x,y) \geqslant k - c$ for appropriate c.

We will define two words x and y satisfying the conditions of the theorem. The words x, y must be such that there is no z such that $K(z) \leqslant k-c$, $I(z:x) \geqslant a$, $I(z:y) \geqslant b$. The last two inequalities means that $K(x|z) \leqslant K(x) - a$, $K(y|z) \leqslant K(y) - b$.

Let c_1 stand for the constant such that $K(u) \leqslant l(u) + c_1$ for all $u \in \Xi$. Let us define

$$
\begin{aligned}
M_1 &= \{(x,y) \in \Xi \times \Xi \mid \exists z \in \Xi\ K(z) \leqslant k-c, K(x|z) \leqslant m + c_1 - a,\\
&\qquad K(y|z) \leqslant n + c_1 - b\},\\
M_2 &= \{(x,y) \mid l(y) = n,\ K(x) < m-2\},\\
M_3 &= \{(x,y) \mid l(x) = m,\ K(y) < n-2\},\\
M_4 &= \{(x,y) \mid K(x,y) < m+n-i-2\}.
\end{aligned}
$$

We claim that for sufficiently large $c \in \mathbb{N}$ the inequality $|M_1 \cup M_2 \cup M_3 \cup M_4| < 2^{m+n}$ holds. To prove this claim we will estimate $|M_1|$, $|M_2|$, $|M_3|$, $|M_4|$. Remember the following well known inequalities $|\{x \in \Xi \mid K(x) \leqslant l\}| < 2^{l+1}$, $\forall y \in \Xi\, |\{x \in \Xi \mid K(x|y) \leqslant l\}| < 2^{l+1}$. These inequalities yield

$$
\begin{aligned}
|M_1| &\leqslant \sum_{z:\, K(z) \leqslant k-c} \left|\{(x,y) \mid K(x|z) \leqslant m + c_1 - a,\ K(y|z) \leqslant n + c_1 - b\}\right|\\
&\leqslant 2^{k-c+1} \cdot 2^{m+c_1-a+1} \cdot 2^{n+c_1-b+1} = 2^{m+n+(k-a-b)+2c_1+3-c}\\
&\leqslant 2^{m+n+2c_1+3-c} \quad (\text{as } k \leqslant a+b),\\
|M_2| &< 2^{m-2} \cdot 2^n = 2^{m+n-2}, \qquad |M_3| < 2^m \cdot 2^{n-2} = 2^{m+n-2},\\
|M_4| &< 2^{m+n-i-2} \leqslant 2^{m+n-2}
\end{aligned}
$$

Therefore, if $c \geqslant 2c_1 + 5$ then $|M_1 \cup M_2 \cup M_3 \cup M_4| < 2^{m+n}$. Hence there is a pair of words x, y such that $l(x) = m$, $l(y) = n$ and $(x,y) \notin M_1 \cup M_2 \cup M_3 \cup M_4$. Let us

take the first pair (x, y) with these properties. We claim that if c is large enough then (x, y) satisfies all requirements.

The conditions $|K(x) - m| \leqslant c$, $|K(y) - n| \leqslant c$ are satisfied because $(x, y) \notin M_2 \cup M_3$ therefore

$$m - 2 \leqslant K(x) \leqslant l(x) + c_1 = m + c_1, \qquad n - 2 \leqslant K(y) \leqslant l(y) + c_1 = n + c_1.$$

As $(x, y) \notin M_1$ there is no z such that $K(z) \leqslant k - c$ and $K(x|z) \leqslant m + c_1 - a$, $K(y|z) \leqslant n + c_1 - b$. We know that $K(x) \leqslant m + c_1$, $K(y) \leqslant n + c_1$. Therefore there is no z such that $K(z) \leqslant k - c$ and $K(x|z) \leqslant K(x) - a$, $K(y|z) \leqslant K(y) - b$.

It remains to prove the difficult assertion

$$|\bar{I}(x : y) - i| \leqslant 5(\log m + \log n) + c.$$

The upper bound of $\bar{I}(x : y)$ is easy: $K(x) \leqslant m + c_1$, $K(y) \leqslant n + c_1$, $K(x, y) \geqslant m + n - i - 2$ (as $(x, y) \notin M_4$) therefore

$$\bar{I}(x : y) \leqslant m + c_1 + n + c_1 - (m + n - i - 2) = i + 2c_1 + 2 \leqslant i + c.$$

Now let us estimate $\bar{I}(x : y)$ from below. We have to estimate $K(x, y)$ by the quantity about $m + n - i$. The main idea is the following: every finite set M can be specified by its cardinality $|M|$ and by an algorithm A which generates all elements of M (and only them). Indeed, given A we can generate the elements of M until we get $|M|$ different elements.

So let us estimate $K(x, y)$. To specify (x, y) it is sufficient to specify m, n and M_1, M_2, M_3, M_4. To specify M_1 it is sufficient to specify the set $\Gamma_1 = \{(p, u) \mid l(p) \leqslant k - c, f_0(p) = u\}$ and the set $\Gamma_2 = \{(p, z, u) \mid g_0(p, z) = u,\ l(p) \leqslant \max(m - a, n - b) + c_1, K(z) \leqslant k - c\}$. To specify M_2, M_3, M_4 it is sufficient to specify $m - 2$, $n - 2$, $m + n - 2$ and the set $\Gamma_3 = \{(p, u) \mid f_0(p) = u,\ l(p) \leqslant m + n - i - 2\}$ (because $m + n - i \geqslant m, n$). To specify Γ_1, Γ_2, Γ_3 it is sufficient to specify $\Gamma = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$ and m, n, a, b, i, c, c_1. Moreover, to specify Γ_1, Γ_2, Γ_3 it is sufficient to specify $|\Gamma|$ and m, n, a, b, i, c, c_1 because given m, n, a, b, i, c, c_1 we can generate the elements of Γ. Therefore there is a specifying method which on the input $\bar{m}\bar{n}\bar{a}\bar{b}\bar{i}\bar{c}\bar{c}_1|\Gamma|$ outputs (x, y). Consequently,

$$\begin{aligned} K(x, y) &\leqslant 2(\log m + \log n + \log a + \log b + \log i) + 2\log c + 2\log c_1 + \log|\Gamma| + c_2 \\ &\leqslant 5(\log m + \log n) + 4\log c + \log|\Gamma| + c_2. \end{aligned}$$

Let us estimate $|\Gamma|$. The number of words of length $\leqslant l$ is less than 2^{l+1}. Therefore

$$\begin{aligned} |\Gamma_2| &< 2^{k-c+1} \cdot 2^{\max(m-a, n-b)+c_1+1} = 2^{\max(k+m-a, k+n-b)-c+c_1+2} \\ &\leqslant 2^{m+n-i-c+c_1+2} \end{aligned}$$

(the last inequality holds because $k \leqslant a + n - i$, $k \leqslant b + m - i$),

$$|\Gamma_3| < 2^{m+n-i-1}, \qquad |\Gamma_1| < 2^{k-c+1} \leqslant 2^{a+n-i-c+1} \leqslant 2^{m+n-i-c+1}.$$

Therefore if $c > c_1 + 1$ then $|\Gamma_1|, |\Gamma_2|, |\Gamma_3| < 2^{m+n-i+1}$ consequently $|\Gamma| < 2^{m+n-i+3}$.

Hence

$$\begin{aligned} K(x,y) &< 5(\log m + \log n) + 4\log c + m + n - i + 3 + c_2, \\ \bar{I}(x:y) &> m - 2 + n - 2 - 5(\log m + \log n) - 4\log c - m - n + i - 3 - c_2 \\ &= i - 5(\log m + \log n) - 4\log c - 7 - c_2. \end{aligned}$$

Let us take $c \in \mathbb{N}$ such that $c > 4\log c + 7 + c_2$. Then we have proved the desired lower bound of $\bar{I}(x:y)$. The theorem is proved.

Let us deduce from Theorem 2 negative solution of A.A.Muchnik's problem. Let us fix some constant c in assertion (1). We will call a pair (x,y) bad if there is no z satisfying (1). We claim that there is bad (x,y) with arbitrary large $K(x)$, $K(y)$ and arbitrary ratios $\bar{I}(x:y)/K(y)$, $\bar{I}(x:y)/K(x)$ belonging to the interval $(0,1)$.

Corollary. *Let $\gamma,\alpha,\beta \in \mathbb{R}$ satisfy inequalities $0 \leqslant \alpha,\beta \leqslant 1$ and $0 < \gamma < \alpha,\beta$. Then for some c for all sufficiently large $j \in \mathbb{N}$ there is a bad pair (x_j,y_j) such that $K(x_j) =_c \alpha \cdot j$, $K(y_j) =_c \beta \cdot j$, $\bar{I}(x_j : y_j) =_{10\log j + c} \gamma \cdot j$.*

Proof. Let us take $\varepsilon > 0$. Define $b = a = (\gamma - \varepsilon)j$, $m = \alpha j$, $n = \beta j$, $i = \gamma j$. Obviously we can choose ε so small that for all j

$$\begin{aligned} (\gamma + \varepsilon)j &\leqslant \min\{a + b, a + n - i, b + m - i\} \\ &= \min\{\gamma - \varepsilon + \gamma - \varepsilon, \gamma - \varepsilon + \beta - \gamma, \gamma - \varepsilon + \alpha - \gamma\} \cdot j. \end{aligned}$$

By Theorem 2 there is a pair (x_j,y_j) such that

$$\begin{aligned} &K(x_j) =_c \alpha j, K(y_j) =_c \beta j, \bar{I}(x_j : y_j) =_{10\log j + c} \gamma j, \\ &C_{(\gamma-\varepsilon)j,(\gamma-\varepsilon)j}(x_j,y_j) \geqslant (\gamma + \varepsilon)j - 10\log j - c. \end{aligned}$$

Evidently for sufficiently large j the pair (x_j,y_j) is bad.

Theorem 3. *There is a constant c such that the following holds. For every $x \in \Xi$ for every $a \leqslant K(x)$ there is $y \in \Xi$ such that*

$$\begin{aligned} &K(y) =_{4\log K(x) + c} K(x), \bar{I}(x:y) =_{6\log K(x) + c} \tfrac{1}{2}K(x), \\ &C_{a,a}(x,y) \geqslant \tfrac{3}{2}a - 4\log K(x) - c. \end{aligned}$$

Proof. Let us fix some $x_0 \in \Xi$ and some $a \in \mathbb{N}$ such that $a \leqslant K(x_0)$. Denote $n = K(x_0)$. It is well known that for every $l \in \mathbb{N}$ there is a prime number p in interval $[l, l+1, \ldots, 2l]$.

Take the least prime $p \geqslant 2^{\lceil \frac{n+1}{2} \rceil}$. As $p \leqslant 2^{\lceil \frac{n+1}{2} \rceil + 1}$, we have $2^{n+1} \leqslant p^2 \leqslant 2^{n+4}$.

Consider the following bipartite graph Γ_1. We will call the vertices from the first part "left vertices" and the vertices from the second part "right vertices".

The set of left vertices is $\mathbb{Z}_p \times \mathbb{Z}_p$, where $\mathbb{Z}_p$ stands for the field of residues modulo p. The set of right vertices is an another copy of $\mathbb{Z}_p \times \mathbb{Z}_p$. Each left vertex $(q,r) \in$

$\mathbb{Z}_p \times \mathbb{Z}_p$ is connected with p right vertices $(q+i, r+qi)$, $i \in \mathbb{Z}_p$. Let us prove that Γ_1 has no closed cycles consisting of four edges (we will call such a cycle "4-cycle").

It is sufficient to prove that if (v_1, v_2) and (v_1, v_3) are edges and $v_2 \neq v_3$ then given v_2, v_3 we can find v_1.

Let $v_1 = (q,r)$, $v_2 = (q+i, r+qi)$, $v_3 = (q+j, r+qj)$. Then given v_2, v_3 we can find $i-j$ and $q(i-j)$. As $\mathbb{Z}_p$ is field we can find q. From q and $q+i$ we can compute i. From $r+qi$ and qi we can compute r.

Furthermore, one can easily verify that for each right vertex v there are exactly p edges incident to v.

Now we will mark left and right vertices with binary words and delete some left vertices. Let us mark ith right vertex with ith word of length $n+4$ (we assume that some computable well ordering on words and vertices is fixed).

We will mark left vertices with the words of entropy $\leqslant n$ as follows. Let us begin to generate the words of entropy $\leqslant n$. The ith generated word marks the ith vertex. As

$$|\{x \in \Xi \mid K(x) \leqslant n\}| < 2^{n+1} \leqslant |\mathbb{Z}_p \times \mathbb{Z}_p|,$$

this is possible. Then we delete all unmarked left vertices. We will denote by Γ_2 the marked graph obtained by this procedure. Note that we can not compute Γ_2 given n because we do not know at which moment the last word x with $K(x) \leqslant n$ is generated. Nevertheless there is a procedure which given n generates all edges of Γ_2 with their markings. We will say "left vertex x" instead of "left vertex with marking x" and for the right vertices also.

Our plan is as follows. We will construct subgraphs M_d of the graph Γ_2 ($d \in \mathbb{N}$). There will exist an algorithm which given n, a, d generates all edges of M_d and only them. Cardinality (i.e. the number of edges) of M_d will be not greater than $2^{\frac{3}{2}n - 4\log n - d + c_2}$ (c_2 is a constant; c_1, as earlier, will be the constant such that $\forall u\, K(u) \leqslant l(u) + c_1$).

Further, M'_d will be the set of those left vertices which are incident to edges only from M_d. Then $|M'_d| \leqslant |M_d|/p \leqslant 2^{n - 4\log n - d + c_2}$. For M'_d, like it holds for M_d, there will be an algorithm generating on n, a, d all vertices of M'_d and only them. In the same way as in the proof of Theorem 2, it will follow from the previous conditions that

$$\begin{aligned} \forall x \in M'_d \; K(x) &\leqslant 2\log n + 2\log a + 2\log d + \log |M'_d| + c_3 \\ &\leqslant 4\log n + 2\log d + n - 4\log n - d + c_4 = n + 2\log d - d + c_4. \end{aligned}$$

Choose d so large that $2\log d - d + c_4 < 0$. Then $x \in M'_d \Longrightarrow K(x) < n$, thus $x_0 \notin M'_d$. Therefore there exists $y_0 \in \Xi$ such that $(x_0, y_0) \in \Gamma_2$ and $(x_0, y_0) \notin M_d$.

Let us turn to defining of M_d.

Let E' be the set of all edges $(x, y) \in \Gamma_2$ such that $K(y) < n - 4\log n - d$. Since each vertex is incident to no more than p edges, we obtain $|E'| \leqslant p 2^{n - 4\log n - d} \leqslant 2^{\frac{n}{2}+2} 2^{n - 4\log n - d} = 2^{\frac{3}{2}n - 4\log n - d + 2}$.

Let E'' be the set of all edges $(x, y) \in \Gamma_2$ such that $K(x, y) < \frac{3}{2}n - 4\log n - d$. This implies that $|E''| < 2^{\frac{3}{2}n - 4\log n - d}$.

For any $z \in \Xi$ we define Γ^z to be the subgraph of Γ_2 consisting of all vertices (left and right) y such that $K(y|z) \leqslant n+4+c_1-a$, and of all edges between these vertices. The left and the right parts of Γ^z consist of less than 2^{n+5+c_1-a} vertices. Let us deduce from this an upper bound of the number of edges of Γ^z. Note that Γ^z has no 4-cycles as a subgraph of Γ_1.

Lemma. *If a bipartite graph has no 4-cycles and the number of vertices in both parts does not exceed m then the number of edges in the graph does not exceed* $m^{3/2}+2m$.

Proof. Let the left vertices of the graph be integers $1,2,\ldots,m$. Let us denote by k_i the number of edges incident to the left vertex i.

Let M_i be the set of all nonordered pairs (j,q), where j and q are right vertices connected with i. Then $|M_i| = \frac{k_i(k_i-1)}{2}$. As the graph has no 4-cycles the sets $M_1,\ldots,M_m$ are pairwise disjoint. As $\left|\bigcup_{i=1}^m M_i\right| \leqslant \frac{1}{2}m(m-1)$, we can conclude that $\sum_{i=1}^m \frac{1}{2}k_i(k_i-1) \leqslant \frac{1}{2}m(m-1)$. Let us deduce from this inequality the inequality $\sum_{i=1}^m k_i \leqslant m^{3/2}+2m$.

Repeat till this is possible the following transformation with the vector $(k_1,\ldots,k_m)$: pick k_i and k_j such that $k_i \leqslant k_j - 2$ then decrease k_j by 1 and increase k_i by 1. After performing this transformation $\sum_{i=1}^m k_i$ is not changed and $\sum_{i=1}^m k_i(k_i-1)$ is not increased because the function $r(r-1)$ is convex. Therefore the inequality

$$\sum_{i=1}^{m} k_i(k_i-1) \leqslant m(m-1) \tag{2}$$

remains valid.

Finally we get a vector consisting (for some k) only of numbers k and $k+1$. Because (2) we have $mk(k-1) \leqslant m(m-1)$; $k(k-1) \leqslant m-1$; $k \leqslant \sqrt{m}+1$.

Hence $\sum_{i=1}^m k_i \leqslant (k+1)m \leqslant m(\sqrt{m}+2)$. □

It follows from the lemma that the number of edges of the graph Γ^z does not exceed $2^{\frac{3}{2}(n+5+c_1-a)} + 2^{n+6+c_1-a} \leqslant 2^{\frac{3}{2}(n-a)+c_5}$.

Let E''' be the set of all edges $(x,y) \in \Gamma_2$ for which there exists $z \in \Xi$ such that $K(z) < \frac{3}{2}a - 4\log n - d$ and $(x,y) \in \Gamma^z$. We have

$$\begin{aligned} |E'''| &\leqslant \sum_{z:\, K(z) < \frac{3}{2}a-4\log n-d} (\text{number of edges of } \Gamma^z) \\ &\leqslant 2^{\frac{3}{2}a-4\log n-d} \cdot 2^{\frac{3}{2}(n-a)+c_5} = 2^{\frac{3}{2}n-4\log n-d+c_5}. \end{aligned}$$

Let $M_d = E' \cup E'' \cup E'''$. First, the existence of algorithms generating the graphs E', E'', E''' given n, a, d is evident. The same statement is true for M_d. Second, $|M_d| \leqslant |E'| + |E''| + |E'''| \leqslant 2^{\frac{3}{2}n-4\log n-d+c_2}$.

Now let us verify that the edge (x_0,y_0) satisfies the conditions of the theorem being proved.

I. $K(y_0) \leqslant l(y_0) + c_1 = n+4+c_1$.

II. $(x_0,y_0) \notin E'$, consequently $K(y_0) \geqslant n - 4\log n - d$.

III. $\bar{I}(x_0 : y_0) = K(x_0) + K(y_0) - K(x_0, y_0)$. As $(x_0, y_0) \notin E''$, we have $K(x_0, y_0) \geqslant \frac{3}{2}n - 4\log n - d$. This implies $\bar{I}(x_0 : y_0) \leqslant n + (n + 4 + c_1) - (\frac{3}{2}n - 4\log n - d) = \frac{n}{2} + 4\log n + c_6$.

IV. The graph Γ_1 is specified by the number n, and the number of its edges is equal to p^3. Therefore, the entropy of each edge does not exceed $2\log n + \log p^3 + c_7$. Thus, $K(x_0, y_0) \leqslant \frac{3}{2}n + 2\log n + c_8$; $\bar{I}(x_0 : y_0) \geqslant n + (n - 4\log n - d) - (\frac{3}{2}n + 2\log n + c_8) = \frac{n}{2} - 6\log n - c_9$.

V. If $K(z) < \frac{3}{2}a - 4\log n - d$ then $(x_0, y_0) \notin E''' \Longrightarrow [K(x_0|z) > n + 4 + c_1 - a$ or $K(y_0|z) > n + 4 + c_1 - a]$. Assume $K(y_0|z) > n + 4 + c_1 - a$, then $I(z : y_0) = K(y_0) - K(y_0|z) < (n + 4 + c_1) - (n + 4 + c_1 - a) = a$. The argument for x_0 is similar. From this it follows that $C_{a,a}(x_0, y_0) \geqslant \frac{3}{2}a - 4\log n - d$.

The theorem is proved.

Let us apply Theorem 3 to arbitrary x and to $a = \frac{2}{5}K(x)$. Obviously, for sufficiently large $K(x)$ for the pair (x, y) there is no z such that (1) holds.

References

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

[2] A.N. Kolmogorov, Three approaches to the quantitative definition of information, Problems in Information Transmission 1 (1) (1965) 1–7.

[3] An.A. Muchnik, On the extraction of common information of two words, In: Pervyi Vsemirnyi Kongress Obshestva matematicheskoi statistiki i teorii veroyatnostei imeni Bernulli. Thesises, vol. 1 M, Nauka, Moscow, 1986, p. 453 (In Russian).

[4] A.K. Zvonkin, L.A. Levin, The complexity of finite objects and the algorithmic concepts of information and randomness, Russ. Math. Surv. 25 (1970) 83–124.