PUBLICATIONS DE L'INSTITUT MATHÉMATIQUE
Nouvelle série, tome 30 (44), 1981, pp. 17–23

ON SOME BASIC PROPERTIES OF THE KOLMOGOROV COMPLEXITY

Dragan Banjević

Abstract. A. N. Kolmogorov in 1964 defined the notion of complexity of a finite word (see [**1,2**]). Some authors defined later some other kinds of complexity (see [**2, 5–13**]). Some basic properties of the Kolmogorov complexity are considered in this paper. Notations, definitions and statements used in this paper are mostly from [**2**].

0. Let us consider the set S of all finite words over $\{0,1\}$. By definition $\Lambda \in S$, Λ-empty word. The lenght of word $x = a_1a_2\dots a_n,\ a_i \in \{0,1\}$ will be denoted by $l(x) = n$, $l(\Lambda) = 0$. In the sequel the following one-to-one correspondence of the set S onto the set $\{0,1,2,\dots\}$ will be made use of:

word	Λ	0	1	00	01	10	11	000	001	010	011 ...
number	0	1	2	3	4	5	6	7	8	9	10 ...

or $x = a_1a_2\dots a_n \leftrightarrow 2^n - 1 + \sum_{i=1}^{n} a_i 2^{n-i}$. For example $x = 00\cdots 0 \leftrightarrow 2^n - 1,\ y = 11\cdots 1 \leftrightarrow 2(2^n - 1),\ l(x) = l(y) = n$. The symbol x will denote both the word and its corresponding number. For two functions F, G on S we write $F \preceq G$ when $(\exists c)(\forall x \in S)(F(x) \le G(x) + c)$ and $F \asymp G$ when $F \preceq G$ and $G \preceq F$. The concatenation of words x and y we denote by xy. One-to-one function $\Phi : S^2 \to S$ is called the numeration of S^2. Denote by $x \circ y = \Phi(x,y)$. For $x = a_1a_2\dots a_n$ let $\bar{x} = a_1a_1a_2a_2\dots a_na_n01$ and $\overline{\Lambda} = 01$. Then $x \circ y = \bar{x}y$ is one numeration of S^2. We have

$$l(\bar{x}y) = 2l(x) + 2 + l(y) \asymp 2l(x) + l(y),$$
$$l(x) \asymp \log_2 x.$$

LEMMA: *There is no numeration such that $l(x \circ y) \preceq l(x) + l(y)$.*

Proof: Let the function Φ be a numeration and $(\exists)(\forall (x,y))(l(x \circ y) \le l(x) + l(y) + c)$. Let

$$S_k = \{(x,y) : l(x) + l(y) + c = k\},\ S'_k = \{x \circ y : l(x \circ y) \le k\}$$

Denote by $|A|$ the number of elements in A. Then

$$|S_k| = 2^{k-c}(k-c+1),\ |S'_k| \le 2^{k+1}-1.$$

For k sufficiently large $|S_k| > |S'_k|$. On the other hand $((x,y)\in S_k) \Rightarrow (x\circ y \in S'_k)$ implying $|S_k| \le |S'_k|$, which is contradiction. ▲

Notice that if $x \circ y = \overline{l(x)}xy$ and $\varepsilon > 0$, then

$$l(x\circ y) \preceq l(x) + l(y) + 2\log l(x) \preceq (1+\varepsilon)l(x) + l(y).$$

1. In what follows all considered functions F, G, H, Φ, ... are partial recursive functions. Following Kolmogorov, we define the complexity of word x with respect to $F^1, F^1 : S \to S$ by

$$K_{F^1}(x) = \min\{l(p) : F^1(p) = x\},$$

where by definition $\min \varnothing = \infty$.

The conditional complexity of x, given y, with respect to F^2, $F^2 : S^2 \to S$ is

$$K_{F^2}(x/y) = \min\{l(p) : F^2(p,y) = x\}.$$

Kolmogorov and Solomonoff proved (see [**2**]) that there exist optimal functions F_0^1, F_0^2 (but not unique) such that for any functions F^1, F^2

$$K_{F_0^1}(x) \preceq K_{F^1}(x),\ K_{F_0^2}(x/y) \preceq K_{F^2}(x/y).$$

The complexity of x with respect to a fixed optimal function $F_0^1(F_0^2)$ we shall call simply the complexity of x and denote by $K(x)(K(x.y))$. We denote by $p_x(p_x^y)$ any program for which $F_0^1(p_x) = x,\ l(p_x) = K(x)(F_0^2(p_x^y, y) = x,\ l(p_x^y) = K(x/y))$. We can define programs p_x and p_x^y uniquele, but those functions of x and y are not recursive in general. On the other hand there is an effective procedure for computing p_x given x, $K(x)$: Use the algorithm for computing F_0^1 and let it to t operations on words p, $l(p) = b$, $t = 1,2,\ldots$ (for details see Remark 0.1. [**2**]). Then we define the recursive function $J(a,b)$ which equals to the first p for which $F_0^1(p) = a$, and then $p_x = J(x, K(x))$. In the same manner we can define p_x^y. In the following we assume $p_x = J(x, K(x))$.

The complexity satisfies some basic properties [**2**]):

(1) $K(x/\Lambda) \asymp K(x),\ K(x/y) \preceq K(x) \preceq l(x)$,

(2) $K(F(x)) \preceq K(x)$,

(3) $\lim\limits_{x\to\infty} K(x) = \infty$,

(4) $|K(x+h/y) - K(x/y)| \preceq 2K(h) \preceq 2l(h)$,

(5) $n \le \max\{K(x) : l(x) = n\} \preceq n$,

$\max\{K(x/l(x)) : l(x) = n \asymp n\}$,

and in general, for arbitrary set N, and arbitrary y

$\max\{K(x/y) : x \in N\} \geq l(|N|) - 1 \asymp \log|N|.$

We give some other properties of the complexity in the following.

(a) $K(p_x^y) \asymp l(p_x^y) = K(x/y),$

$K(p_x^y) \preceq K(p),\ p$ such that $F_0^2(p, y) = x.$ ▲

(b) If $F(x) \asymp G(x)$, then $K(F(x)/y) \asymp K(G(x)/y)$, $K(y/F(x)) \asymp K(y/G(x))$.

Proof: We can prove (b) using (e) in the following. F and G are not necessarily recursive. ▲

(c) $K(F(x)/y) \preceq K(x/G(y)),$

$K(F(x,y)/y) \preceq K(x/y).$ ▲

If, for fixed y, F is one-to-one function of x, then $K(F(x,y)/y) \preceq K(x/y)$. For example $K(xy/y) \asymp K(x/y)$.

(d) $K(x/y) \asymp K(|x-y|/y) \preceq K(|x-y|),$

$K(x/p_x^y) \preceq K(y).$ ▲

(e) $|K(x+h/y+l) - K(x/y) \preceq 2K(\bar{h}l),$

$I(y:x) = K(x) - K(x/y) \preceq 2K(y).$ ▲

(f) For any numeration and any F

$K(F(x,y) \preceq K(x \circ y).$ ▲

For example $K(x \circ y) \asymp K(\bar{x}y)$, and $\max\{K(x), K(y)\} \asymp K(x \circ y)$.

Remark 1.1. $K(\bar{x}y \preceq K(x) + K(y)$ is not valid but $K(\bar{x}y) \preceq K(x) + K(y) + 2\log\min\{K(x), K(y)\}$. We see that the function $x \circ y = P\bar{x}y$ is a numeration of S^2 and $l(x \circ y) = K(\bar{x}y)$, which implies (in view of Lemma in 0.) that $K(\bar{x}y) \preceq l(x) + l(y)$ is not true, and then $K(\bar{x}y) \preceq K(x) + K(y)$ does not hold neather. On the other hand $K(x) + K(y) \preceq K(\bar{x}y)$ is not valid because in the opposite case for $x = y$ we have $2K(x) \preceq K(\bar{x}x) \preceq K(x)$, or $K(x) \preceq 0$. For the proof of the second inequality it is sufficient to consider programs $\overline{l(p_x)}p_xp_y$ and $\overline{l(p_y)}p_yp_x$ with respect to specified function G. It is interesting that $K(xy) \preceq l(x) + l(y)$ but $K(xy) \preceq K(x) + K(y)$ does not hold(consider previous remark and $K(\bar{x}) \asymp K(x)$). ▲

(g) $\Pi(n) = \min\{K(x) : l(x) = n\} \asymp K(n) \preceq \log n.$

Proof: Let $K(l(x)) \leq K(x) + c_1,\ K(2^x - 1) \leq K(x) + c_2$, folowing (2). Then $K(n) \leq \Pi(n) + c_1 \leq K(2^n - 1) + c_1 \leq K(n) + c_1 + c_2$. ▲

(h) Let $A_m = \{x : K(x) \leq m\}$, $B_m = \{x : K(x/m) \leq m\}$. Then: (i) $m - 2\log m \preceq \log|A_m| \preceq m$, (ii) $\log|B_m| \asymp m$.

Proof: (i) We have immediately $|A_m| \leq 2^{m+1} - 1$. Let for given $p = \bar{a}b$, $x = G(p)$ be such that: We choose the set A of exactly b words y such that $K(y) \leq a$ (see Theorem 1.6) in [**2**]), and $x = G(p)$ is the first y such that $y \notin A$. Then, if $a = m$. $b = |A_m|$, we have $x = G(p) \notin A_m$, and

$m < K(x) \preceq K_G(x) \leq l(p) \preceq \log |A_m| + 2\log m.$ ▲

(j) $\lim_{y\to\infty} K(x/y) \preceq 0$ is not true (compare (3)), but

$\lim_{y\to\infty} \inf K(x/y) \preceq 0.$

Proof: Let $(\exists c)(\forall x)(\exists y_0)(\forall y \geq y_o)(K(x/y) \leq c)$. Then $l(p_x^y) \leq c$ and the number of such programs is at most $N = 2^{c+1} - 1$. Let $M > N$ and consider $0 < x_1 < x_2 < \cdots < x_M$. Let y_i is chosen such that for $y \geq y_i$, $K(x/y) \leq c$. Let $y_0 = \max\{y_1, y_2, \cdots, y_M\}$, then for arbitrary $y \geq y_0$, $F_0^2(p_{x_i}^y, y) = x_i$, $i = 1, 2, \cdots, M$. and programs $p_{x_i}^y$ are all different, which is a contradiction. It means that $(\forall c)(\exists x)(\forall y_0)(\exists y \geq y_0)(K(x/y) \geq c)$. It is easy to see that $K(x/\bar{x}i) \preceq 0$ for all i, or $\lim_{y\to\infty} \inf K(x/y) \preceq 0$. ▲

2. The complexity of a sequence of words $x_1, x_2, \dots, x_m$, given a sequence $y_1, y_2, \dots, y_k$, with respect to a sequence of functions $F = (F_1, F_2, \dots, F_m)$, $F_i : S^{k+1} \to S$, can be defined as

$$K_F(x_1, \dots, x_m/y_1, \dots, y_k) = \min\{l(p) : F_i(p, y_1, \dots, y_k) = x_i,\ i = 1, 2, \dots, m\}.$$

It can be shown that there exists an optimal sequence $F_0 = (F_{01}, \dots, F_{0m})$ such that $K_{F_0} \preceq K_F$, and we define $K(x_1, \dots, x_m/y_1, \dots, y_k)$ as the complexity with respect to F_0. In the similar way we can define $K(x_1, \dots, x_m)$. It can be proved that

(*) $$K(x_1, \dots, x_m/y_1, \dots, y_k) \asymp K(x_1 \circ \cdots \circ x_m/y_1, \dots, y_k),$$

(**) $$K(x/y_1, \dots, y_k) \asymp K(x/y_1 \circ \cdots \circ y_k),$$

where $z_1 \circ \cdots \circ z_j$ is notation for a numeration of S^j. Considering $(*)$ and $(**)$ we have $K(x, y) \asymp K(\bar{x}y)$, $K(x/y, z) \asymp K(x/\bar{y}z)$. Some authors define directly $K(x, y) = K(\bar{x}y)$ (see [**5**] p. 332, for example). It is easy to show some properties of the complexity, for example

$$K(x, F(x)) \asymp K(x),\ K(x/y, x) \preceq K(x/F(y, z)),$$
$$K(F(x, y)) \preceq K(x, y),\ K(F(x), G(y)) \preceq K(x, y),$$
$$K(x/y), z) \preceq K(x/F(y),\ G(z)),$$
$$|K(x + h,\ y + l) - K(x, y)| \preceq 2K(h, l),$$
$$K(x, y/l(x), l(y)) \preceq K(x, y/l(x)) \preceq l(x) + K(y/x) \preceq l(x) + l(y),$$
$$\max\{K(x, y/l(x),\ l(y)) : l(x) = n,\ l(y) = m\} \asymp n + m,$$
$$\max\{K(x/l(x),\ s(x)) : l(x) = n,\ s(x) = s\} \asymp \log \binom{n}{s},$$

where for $x = a_1 a_2 \dots a_n$, $s(x) = \sum_{i=1}^{n} a_i$ (it is an immediate consequence of 1.(5), but see [**1, 3, 4**]).

We give some other properties.

(a) $K(x,y) \preceq K(p_x,y) \asymp K(x,y,K(x)) \preceq$
$\preceq K(p_x,p_y) \asymp K(x,y,K(x),\ K(y)).$ ▲

(b) $K(y/x,\ K(x)) \asymp K(y/p_x) \preceq K(y/x).$ ▲

(c) $K(x/z) \preceq K(x/y,\ K(y/z)) + 2K(y/z),$
$(K(x/z) \preceq 2K(x/y,\ K(y/z)) + K(y/z),$ ▲

(d) $K(x,y) \preceq K(x) + 2K(y/x),\ K(x)),$
$K(x,y/K(x)) \preceq K(x) + K(y/x,\ K(x)).$ ▲

If we put $z = \Lambda$ in (c) we have

$$-2K(y/x) \preceq -2K(y/p_x) \preceq K(x) - K(y) \preceq 2K(x/py) \preceq 2K(x/y).$$

Remark 2.1. Theorem 1. (Levin) in [**12**] states that $KP(x,y) \asymp KP(x) + KP(y/x, KP(x))$ (also see Th. 5.1. (b) in [**5**]), where $KP(x)$ is some variant of complexity (see [**5, 10-13**]). But for the Kolmogorov complexity $K(x,y) \preceq K(x) + K(y/x,\ K(x))$ is not valid. We shall prove $(\forall c)(\exists(x,y))(K(x) + K(y/p_x) \geq K(\bar{x}y) + c)$. Following 1. (5) $(\forall l_0)(\exists x)(l(x) = l_0,\ \ K(x/l_0) \geq l_0 - 1)$. In view of 1. (1) and 1. (2), $(\exists c_1)(\forall x)(l(x) \geq K(x) - c_1)$, $(\exists c_2)(\forall x)$ $(K(x) \geq K\overline{l(x)}x) - c_2)$ and following 2. (a) $(\exists c_3)(\forall(x,y))(K(\bar{p}_x y) \geq K(\bar{x}y) - c_3)$.

Let x be chosen such that for fixed $c, l(p_x) = K(x) \geq c + c_1 + c_2 + c_3 + 1$. Let y be chosen such that $l(y) = l_0 = p_x$, and $K(y/l_0) \geq l_0 - 1 = l(y) - 1$. Then $K(y/p_x) \geq l(y) - 1 \geq K(y) - c_1 - 1 \geq K(\overline{l(y)}y) - c_2 - c_1 - 1 = K(\bar{p}_x y) - c_2 - c_1 - 1 \geq K(\bar{x}y) - c_3 - c_2 - c_1 - 1$ and $K(x) + K(y/p_x) \geq K(\bar{x}y) + c$. ▲

(e)

(i) $\min\{K(p_x/x) : l(x) = n\} \preceq 0,$

(ii) $\log n - \log\log n \preceq \max\{K(p_x/x) : l(x) = n\} \preceq \log n,$

(see Theorem 2. in [**12**] and Theorem 5.1. (f) in [**5**].)

Proof: Basic ideas for proving follow the proof of Theorem 2. in [**12**]. We have

$$K(p_x/x) \asymp K(K(x)/x) \preceq l(K(x)) \preceq l[l(x)].$$

(i) Let $K(x) \leq l(x) + c$, and $A = \{x : |K(x) - l(x)| \leq c\}$.

Following 1.(5), we have $(\forall n)(\exists x)(l(x) = n,\ x = A)$. Then by 1. (c) and 1. (d), for $x \in A$, $K(K(x)/x) \preceq K(K(x)/l(x)) \preceq l(|K(x) - l(x)|) \preceq l(c)$, or $\min\{K(p_x/x) : l(x) = n\} \preceq 0$.

(ii) Let $r = r(n) = \max\{K(p_x/x) : l(x) = n\}$. Then $r \preceq \log n$, and $(\forall x, l(x) = n)(\exists p)(l(p) \leq r,\ F_0^2(p,x) = p_x)$. Let $M_i = \{x : l(x) = n$ and for at least i programs $p, l(p) \leq r,\ F_0^1(F_0^2(p,x)) = x\}$, $i = 1, 2, \ldots$. Then $|M_1| = 2^n$, $M_1 \supset M_2 \supset \cdots \supset M_j \supset M_{j+1}$, $M_j \neq \varnothing$, $M_{j+1} = \varnothing$, and $2^{r+1} - 1 \geq j$ or $r \preceq \log j$.

We shall prove by induction that $\log|M_i| \geq n - (i-1)(3\log n + k)$, $i = 1.2, \ldots, j$, where k—constant. For $i = 1$, the proposition is valid.

Let the function G be defined for programs p of the form $p = \overline{l(a)l(b)l(c)l(d)}$ $abcde$, in the following way:

I Let the algorithm for computing $F_0^1(F_0^2(p,x))$ do t operations on words $x, l(x) = a$, and programs $p, l(p) \leq b$ (see Remark 0.1. in [**2**]), $t = 1, 2, \ldots$. We stop the computation when we get exactly e words x such that for at least $c+1$ programs p $F_0^1(F_0^2(p,x)) = x$.

II From the set of the remaing $2^a - e$ words x, we take the first word x such that for exactly c programs $p\, F_0^1(F_0^2(p,x)) = x$, and $\min\{l[F_0^2(p,x)]\} \geq \log[2^{a-(c-1)(3\log a+\varphi(d))} - e] - 2$, where $\varphi(d)d = d + \log d + 2\log\log d + B$, B an absolute constant.

Now, let $K(x) \leq K_G(x) + A$. Suppose that $\log|M_i| \geq n - (i-1)(3\log n + \varphi(A)) = m_i$. Then $|M_i - M_{i+1}| = |M_i| - |M_{i+1}| \geq 2^{m_i} - |M_{i+1}|$. If $2^{m_i} - |M_{i+1}| \leq 0$ is true, then $|M_{i+1}| \geq 2^{m_i+1}$. Supose that $2^{m_i} - |M_{i+1}| > 0$. Then in $M_i - M_{i+1}$ there exists x such that $K(x) \geq \log[2^{m_i} - |M_{i+1}|] - 2$, and we can get such x if we compute $G(p)$ for $a = n,\ b = r,\ c = i,\ d = A,\ e = |M_{i+1}|$. Then $\log[2^{m_i} - |M_{i+1}|] - 2 \leq K(x) \leq K_G(x) + A \leq l(p) + A \leq \log|M_{i+1}| + 3\log n + A + \log A + 2\log\log A + D$, and $\log|M_{i+1}| \geq n - i(3\log n + \varphi(A)),\ B = D + 3$.

In the same manner we have $2^{m_{j+1}} \preceq 0$ or $j \succeq \frac{n}{3\log n + \varphi A}$ and $r \succeq \log j \succeq \log n - \log\log n$. ▲

For example, using (e), we have $(\forall c)(\exists(x,y))(K(y/x) \geq (K(y/p_x) + c)$ (put $y = p_x$).

REFERENCES

[1] Kolmogorov A. N., *Logical basic for information theory and probability theory*, IEEE, Trans. of Inf. Th., vol. IT-14, **5**, 1968, 662–664.

[2] Звонкин, А. К., Левин, Л. А., *Сложность конечных обьектов и обоснование понятий информации и случайности с помощью теории алгоритмов*, УМН, XXV, **6**, 1970, 85–127.

[3] Martin-Lof, P., *The definition of random sequence*, Inf. and Con., **9**, 1966, 602–619.

[4] Fine, T. L., *On the apparent convergence of relative frequence and its inplications*, IEEE, Trans. of Inf. Th., vol. IT-16, **3**, 1970, 251–257.

[5] Chaitin, G. J., *A theory of program size formally identical to information theory*, JACM, 22 **3**, 1975, 329–340.

[6] Chaitin, G. J., *Program size, oracles, and the jump operations*, Osaka J. Math. **14**, 1977, 139–149.

[7] Loveland, D. V., *A variant of the Kolmogorov concept of complexity*, Inf. and Con., **15**, 1969, 410–526.

[8] Schnorr, C. P., *Process complexity and effective random tests*, J. of Comp. and Syst. Sci. **7**, 1973, 376–388.

[9] Левин, Л. А., *О понятии случайной последовательности*, ДАН СССР, 1973, 212, **3**, 548–550.

[10] Левин, Л. А., *Законы сохранения (невозрастения) информации и вопросы обоснования теории вероятностей*, ППИ, X, 1974, **3**, 30–35.

[11] Левин, Л. А., *О различных мерах сложности конечних объектов (аксиоматическое описание)*, ДАН СССР, 1976, 227, **4**, 804–807.

[12] Гач, П., *О симетрии алгоритмической информации*, ДА, СССР, 1974, 218, **6**, 1265–1267.

[13] Gacs, P., *Exact expressions for some randomness tests*, Z. Math. Log. Grund. Math., 26, **5**, 1980, 385–394.

Institut za matematiku
11000 Beograd, Studentski trg 16

Corrections of some relevant printing errors in paper "Algorithmical definition of finite Markov sequence", D. Banjević and Z. Ivković, Publ. l'Inst. Math. **28**(**42**), 1980. pp. 13–17.

Page	Printed:	Correction:
13^{11}	$H_i(x_0, t_{x1}; \dots; x_1, t_{xi})$	$H_i(x_1, t_{x_1}; \dots; x_i, t_{xi})$
14^{7}	t_{i_j-1}	t_{i_j-1}
14^{8}	$\frac{1}{v_i}\sum_{j-u}^{k}(1-t_{i_j-1})1-t_{i_j})$	$\frac{1}{v_i}\sum_{j-u}^{k}(1-t_{i_j-1})(1-t_{i_j})$
14_{6}	$R_3, \dots\}$	$R_2, \dots\}$
15^{10}	$V = V_0$	$v = v_0$
15^{16}	(n, δ, p)	(m, δ, p)
15_{2}	$\geq P(v_0 \geq n_0, \Delta_0 \geq \varepsilon_0)+$	$\leq P(v_0 \geq n_0, \Delta_0 \geq \varepsilon_0)+$
16^{2}	$i_1, i_3 \dots,$	$i_1, i_2, \dots,$
16^{12}	$v_+^0 = n_0$	$v_0^* = n_0$
16^{13}	$\sum_{i=0}^{j}$	$\sum_{i=1}^{j}$
16_{12}	$\frac{1}{V_0^*}$	$\frac{1}{v_0^*}$
16_{8}	$\leq P[$	$\leq \varrho[$
16_{6}	sequence is $1 - P(\mathcal{R}) < 0$	sequence is $1 - P(\mathcal{R}) > 0$
17^{1}	REFRENCEES	REFERENCES
17_{3}	$538 - 550$	$548 - 550$