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 \( \Lambda \in S \), \( \Lambda \)-empty word. The length of word \( x = a_1 a_2 \ldots a_n \), \( a_i \in \{0, 1\} \) 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 \( \{0, 1, 2, \ldots\} \) will be made use of:

<table>
<thead>
<tr>
<th>word</th>
<th>( \Lambda )</th>
<th>0</th>
<th>1</th>
<th>00</th>
<th>01</th>
<th>10</th>
<th>11</th>
<th>000</th>
<th>001</th>
<th>010</th>
<th>011</th>
<th>\ldots</th>
</tr>
</thead>
<tbody>
<tr>
<td>number</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>\ldots</td>
</tr>
</tbody>
</table>

or \( x = a_1 a_2 \ldots a_n \leftrightarrow 2^n - 1 + \sum_{i=1}^{n} a_i 2^{n-i} \). For example \( x = 00 \ldots 0 \leftrightarrow 2^n - 1 \), \( y = 11 \ldots 1 \leftrightarrow 2(2^n - 1) \), \( l(x) = l(y) = n \). 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 \( (\exists c)(\forall x \in S)(F(x) \leq G(x) + c) \) 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 \( x = a_1 a_2 \ldots a_n \) let \( \bar{x} = a_1 a_2 a_3 \ldots a_n a_1 \) and \( \bar{\Lambda} = 01 \). Then \( x \circ y = \bar{xy} \) is one numeration of \( S^2 \).

We have

\[
\begin{align*}
l(\bar{xy}) &= 2l(x) + 2 + l(y) \approx 2l(x) + l(y), \\
l(x) &\asymp \log_2 x.
\end{align*}
\]

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

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

\[
S_k = \{(x, y) : l(x) + l(y) + c = k\}, \quad S'_k = \{x \circ y : l(x \circ y) \leq k\}
\]
Denote by $|A|$ the number of elements in $A$. Then

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

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 $|S_k| \leq |S'_k|$, which is contradiction. ▲

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

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

1. In what follows all considered functions $F, G, H, \Phi, \ldots$ are partial recursive functions. Following Kohnogorov, we define the complexity of word $x$ with respect to $F^1, F^2 : S \to S$ by

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

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

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 $F^1_0, F^2_0$ (but not unique) such that for any functions $F^1, F^2$

$$K_{F^1}(x) \preceq K_{F^1_0}(x), \quad K_{F^2}(x/y) \preceq K_{F^2_0}(x/y).$$

The complexity of $x$ with respect to a fixed optimal function $F^1_0$ ($F^2_0$) we shall call simply the complexity of $x$ and denote by $K(x)(K(x, y))$. We denote by $p_x(p^y_x)$ any program for which $F^1_0(p_x) = x$, $l(p_x) = K(x)(F^1_0(p^y_x), y) = x$, $l(p^y_x) = K(x, y)$. We can define programs $p_x$ and $p^y_x$ unique, 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 $F^1_0$ 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 $F^1_0(p) = a$, and then $p_x = J(x, K(x))$. In the same manner we can define $p^y_x$. In the following we assume $p_x = J(x, K(x))$.

The complexity satisfies some basic properties [2]:

1. $K(x \Lambda) \preceq K(x)$, $K(x/y) \preceq K(x) \preceq l(x)$,
2. $K(F(x)) \preceq K(x)$,
3. $\lim_{x \to \infty} K(x) = \infty$,
4. $|K(x + h/y) - K(x/y)| \preceq 2K(h) \preceq 2l(h)$,
5. $n \leq \max\{K(x) : l(x) = n\} \preceq n$,
6. $\max\{K(x/l(x)) : l(x) = n \times n\}$,

and in general, for arbitrary set $N$, and arbitrary $y$
max\{K(x/y) : x \in N\} \geq l(|N|) - 1 \simeq \log |N|.

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

(a) \(K(p_x) \simeq l(p_x) = K(x/y)\),
\[ K(p_x) \preceq K(p) \] such that \(F_0^p(p, y) = x\).

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

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

(c) \(K(F(x)/y) \preceq K(x/G(y))\),
\[ K(F(x), y) \preceq K(x/y) \] if, for fixed \(y\), \(F\) is one-to-one function of \(x\), then \(K(F(x), y) \preceq K(x/y)\).

For example \(K(xy/y) \simeq K(x/y)\).

(d) \(K(x/y) \simeq K(|x - y|/y) \preceq K(|x - y|)\),
\[ K(x/p_x) \preceq K(y) \] if \(K(x + h/y + l) - K(x/y) \preceq 2K(hl)\),
\[ I(y : x) = K(x) - K(x/y) \preceq 2K(y) \]
\[(e)\]
For any enumeration and any \(F\)
\[ K(F(x, y) \preceq K(x \circ y) \]
For example \(K(x \circ y) \simeq K(\bar{x}y)\), and \(\max \{K(x), K(y)\} \simeq K(x \circ y)\).

\[ K(\bar{x}y) \preceq K(x) + K(y) \]
\[ K(\bar{x}y) \preceq K(x) + K(y) \]
\[ \Pi(n) = \min \{K(x) : l(x) = n\} \preceq K(n) \preceq n \]
\[ \Pi(n) = \min \{K(x) : l(x) = n\} \preceq K(n) \preceq n \]
\[ \Pi(n) = \min \{K(x) : l(x) = n\} \preceq K(n) \preceq n \]
\[ \Pi(n) = \min \{K(x) : l(x) = n\} \preceq K(n) \preceq n \]
\[ \Pi(n) = \min \{K(x) : l(x) = n\} \preceq K(n) \preceq n \]
\[ \Pi(n) = \min \{K(x) : l(x) = n\} \preceq K(n) \preceq n \]
\[ \Pi(n) = \min \{K(x) : l(x) = n\} \preceq K(n) \preceq n \]
\[ \Pi(n) = \min \{K(x) : l(x) = n\} \preceq K(n) \preceq n \]
\[ \Pi(n) = \min \{K(x) : l(x) = n\} \preceq K(n) \preceq n \]
\[ \Pi(n) = \min \{K(x) : l(x) = n\} \preceq K(n) \preceq n \]
\[ \Pi(n) = \min \{K(x) : l(x) = n\} \preceq K(n) \preceq n \]
\[ \Pi(n) = \min \{K(x) : l(x) = n\} \preceq K(n) \preceq n \]
\[ \Pi(n) = \min \{K(x) : l(x) = n\} \preceq K(n) \preceq n \]
\[ \Pi(n) = \min \{K(x) : l(x) = n\} \preceq K(n) \preceq n \]
\[ \Pi(n) = \min \{K(x) : l(x) = n\} \preceq K(n) \preceq n \]
\[ \Pi(n) = \min \{K(x) : l(x) = n\} \preceq K(n) \preceq n \]
\[ m < K(x) \leq K_G(x) \leq l(p) \leq \log |A_m| + 2 \log m. \]

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

\[ \lim_{y \to \infty} \inf K(x/y) \leq 0. \]

Proof: Let \((\exists c)(\forall x)(\exists y_0)(\forall y \geq y_0)(K(x/y) \leq c).\) Then \(l(p^y) \leq c\) 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 \(y_i\) be 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 \(y \geq y_0,\) \(F_0^x(p^y_{y_i}, y) = x_i, i = 1, 2, \cdots, M.\) and programs \(p^y_{y_i}\) are all different, which is a contradiction. It means that \((\forall c)(\exists x)(\exists y_0)(\exists y \geq y_0)(K(x/y) \geq c).\) It is easy to see that \(K(x/x_i) \leq 0\) for all \(i,\) or \(\lim_{y \to \infty} \inf K(x/y) \leq 0.\)

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

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

It can be shown that there exists an optimal sequence \(F_0 = (F_0^1, \ldots, F_0^m)\) such that \(K_{F_0} \leq K_F,\) and we define \(K(x_1, \ldots, x_m/y_1, \ldots, y_k)\) as the complexity with respect to \(F_0.\) In the similar way we can define \(K(x_1, \ldots, x_m).\) It can be proved that

\[(*) \quad K(x_1, \ldots, x_m/y_1, \ldots, y_k) \asymp K(x_1 \circ \cdots \circ x_m/y_1, \ldots, y_k), \]

\[(**) \quad K(x/y_1, \ldots, y_k) \asymp K(x/y_1 \circ \cdots \circ y_k), \]

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

\[
K(x, F(x)) \asymp K(x), \quad K(x/y, x) \asymp K(x/F(y, z)),
\]
\[
K(F(x, y)) \asymp K(x, y), \quad K(F(x), G(y)) \asymp K(x, y),
\]
\[
K(x/y, z) \asymp K(x/F(y), G(z)),
\]
\[
|K(x + h, y + l) - K(x, y)| \leq 2K(h, l),
\]
\[
K(x, y/l(x), l(y)) \asymp K(x, y/l(x)) \asymp l(x) + K(y/x) \asymp l(x) + l(y),
\]
\[
\max \{K(x, y/l(x), l(y)) : l(x) = n, l(y) = m\} \asymp n + m,
\]
\[
\max \{K(x/l(x), s(x)) : l(x) = n, s(x) = s\} \asymp \log \left(\frac{n}{s}\right),
\]

where for \(x = a_1a_2 \ldots a_m, s(x) = \sum_{i=1}^{n} a_i\) (it is an immediate consequence of 1,(5),

but see [1, 3, 4]).

We give some other properties.
On some basic properties of the Kolmogorov complexity

(a) \( K(x, y) \leq K(p_x, y) \preceq K(x, y, K(x)) \preceq K(x, y) \).
(b) \( K(y/x, K(x)) \preceq K(y/p_x) \preceq K(y/x) \).
(c) \( K(x/z) \leq K(x/y, K(y/z)) + 2K(y/z) \),
\( (K(x/z) \leq 2K(x/y, K(y/z)) + K(y/z) \).
(d) \( K(x, y) \leq K(x) + 2K(y/x, K(x)) \),
\( K(x, y/K(x)) \preceq K(x) + K(y/x, K(x)). \)

If we put \( z = A \) in (c) we have
\[
-2K(y/x) \leq -2K(y/p_x) \preceq K(x) - K(y) \preceq 2K(x/p_y) \preceq 2K(x/y).
\]

**Remark 2.1.** Theorem 1. (Levin) in [12] states that \( KP(x, y) \preceq KP(x) + KP(y/x, KP(x)) \) (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 (\( \forall \epsilon(\exists x,y)(K(x)+K(y/p_x) \geq K(\tilde{y}) + c \)). Following 1. (5) \( (\forall l)(\exists x)(l(x) = l_0, K(x/l_0) \geq l_0 - 1) \). In view of 1. (1) and 1. (2), \( (\exists c_1)(\forall l)(l(x) \geq K(x) - c_1, (\exists c_2)(K(x) \geq K(l(x) - c_2) \) and following 2. (a) \( (\exists c_3)(\forall l)(K(\tilde{y}) \geq K(\tilde{y}) - c_3) \).

Let \( x \) be chosen such that for fixed \( c_1, l(p_x) = K(x) > c + c_1 + c_2 + c_3 + 1 \). 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 \( K(y/p_x) \geq l(y) - 1 \geq K(y) - c_1 - 1 \geq K(l(y)) - c_2 - c_1 - 1 \geq K(\tilde{y}) - c_3 - c_2 - c_1 - 1 \) and \( K(x) + K(y/p_x) \geq K(\tilde{y}) + c \).

(e)
(i) \( \min\{K(p_x) : l(x) = n\} \leq 0 \).
(ii) \( \log n - \log \log n \leq \max\{K(p_x) : l(x) = n\} \leq \log n \).

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

(i) \( \min\{K(p_x) : l(x) = n\} \leq c \).

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

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

We shall prove by induction that \( \log |M_i| \geq n - (i - 1)(\log n + k) \), \( k = 1, 2, \ldots \).
Let the function \( G \) be defined for programs \( p \) of the form \( p = \overline{l(a)}l(b)c \) for \( a,b,c \) in \( \{a,b,c\} \), 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 \( p F_0^1(F_0^2(p,x)) = x \).

\[ \mathbb{I} \text{ I } \]

II From the set of the remaining \( 2^a - e \) 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 \( \min \{|F_0^2(p,x)|\} \geq \log(2^{n-e}) = e - 2 \), where \( \varphi(d) = d + \log d + 2 \log \log d + B \), \( B \) 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 \( |M_{i+1}| \geq 2^{m_i+1} \). Suppose that \( 2^{m_i} - |M_{i+1}| \geq 3 \log n + \varphi(A) \). Then \( |M_i - M_{i+1}| \geq 3 \log n + A \log A + 2 \log \log A + D \), and \( \log |M_{i+1}| \geq n - i(3 \log n + \varphi(A)) \), \( B = D + 3 \).

In the same manner we have \( 2^{m_{i+1}} \leq 0 \) or \( j \geq \frac{n}{3 \log n + \varphi(A)} \) and \( r \geq \log j \geq \log n - \log \log n \).

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

REFERENCES


<table>
<thead>
<tr>
<th>Page</th>
<th>Printed:</th>
<th>Correction:</th>
</tr>
</thead>
<tbody>
<tr>
<td>13\textsuperscript{11}</td>
<td>( H_i(x_0, t_{x_1}; \ldots ; x_i, t_{x_i}) )</td>
<td>( H_i(x_1, t_{x_1}; \ldots ; x_i, t_{x_i}) )</td>
</tr>
<tr>
<td>14\textsuperscript{9}</td>
<td>( t_{i_{j-1}} )</td>
<td>( t_{i_{j-1}} )</td>
</tr>
<tr>
<td>14\textsuperscript{9}</td>
<td>( \frac{1}{k} \sum_{j=1}^{k} (1 - t_{i_{j-1}}) (1 - t_{i_j}) )</td>
<td>( \frac{1}{k} \sum_{j=1}^{k} (1 - t_{i_{j-1}}) (1 - t_{i_j}) )</td>
</tr>
<tr>
<td>14\textsuperscript{8}</td>
<td>( R_3, \ldots { )</td>
<td>( R_2, \ldots { )</td>
</tr>
<tr>
<td>15\textsuperscript{10}</td>
<td>( V = V_0 )</td>
<td>( v = v_0 )</td>
</tr>
<tr>
<td>15\textsuperscript{10}</td>
<td>( (\nu, \delta, p) )</td>
<td>( (\nu, \delta, p) )</td>
</tr>
<tr>
<td>15\textsuperscript{2}</td>
<td>( \geq P(\nu_0 \geq n_0, \Delta_0 \geq \varepsilon_0) + )</td>
<td>( \leq P(\nu_0 \geq n_0, \Delta_0 \geq \varepsilon_0) + )</td>
</tr>
<tr>
<td>16\textsuperscript{2}</td>
<td>( i_1, i_2, \ldots , )</td>
<td>( i_1, i_2, \ldots , )</td>
</tr>
<tr>
<td>16\textsuperscript{12}</td>
<td>( v_0^* = n_0 )</td>
<td>( v_0^* = n_0 )</td>
</tr>
<tr>
<td>16\textsuperscript{13}</td>
<td>( \sum_{i=0}^{j} )</td>
<td>( \sum_{i=1}^{j} )</td>
</tr>
<tr>
<td>16\textsuperscript{12}</td>
<td>( \frac{1}{v_0^{*}} \leq P[ )</td>
<td>( \frac{1}{v_0^{*}} \leq P[ )</td>
</tr>
<tr>
<td>16\textsuperscript{8}</td>
<td>sequence is ( 1 - P(\mathcal{R}) &lt; 0 )</td>
<td>sequence is ( 1 - P(\mathcal{R}) &gt; 0 )</td>
</tr>
<tr>
<td>17\textsuperscript{1}</td>
<td>REFERENCES</td>
<td>REFERENCES</td>
</tr>
<tr>
<td>17\textsubscript{3}</td>
<td>538 – 550</td>
<td>548 – 550</td>
</tr>
</tbody>
</table>