Constructing of sequences, with complex finite subsequences

Andrei Yu. Rumyantsev

Moscow State Lomonosov University

Mathematics Department

Logic and Algorithms Theory Division

Levin Lemma

Lemma (Levin). Assume that $\alpha < 1$. Then there exists an infinite sequence of 0 and 1 such that each its substring of (sufficiently large) length n has complexity at least αn.

1. Substrings with small complexity are forbidden.

For example all sufficiently long squares are forbidden.

2. A random sequence does not satisfy the property.

3. The bound αn can not be improved. If the sequence does not contain some substrings then complexities of all substrings are less than $(1-\varepsilon)n$.

The main theorem

Theorem (Main). Let $0 < \gamma < 1$. Then there exist a binary sequence ω and an integer c such that for any finite set A of cardinality at least c the inequality

$$\mathrm{K}(A, \omega(A) \mid t) \geq \gamma |A|$$

holds for some $t \in A$.

$$\omega = 0\mathbf{1}0\mathbf{0}11\mathbf{0}101\ldots$$

$$A = \{1, 3, 6\},\ \omega(A) = 100$$

Equivalent formula: $\mathrm{K}(\omega(A) \mid A) \geq \gamma |A| - \max\limits_{t \in A} \mathrm{K}(A \mid t)$

($\omega(A)$ has high complexity relativly A for “good” finite sets A.)

Consequences

- There exists a two dimensional sequence with complex rectangles (for any $\alpha < 1$ there exists a two dimensional sequence such that each its rectangle of area n has complexity at least αn).
- (For any $\alpha < 1$) there exists a sequence such that all its computable permutations satisfy the condition of Levin Lemma.

More involved applications of the main result

Application 1. Almost periodic sequences

Definition.

A sequence ω is almost periodic if each its substring occurs in ω infinitely many times with bounded intervals between neighbouring occurences.

Theorem. For any $\alpha < 1$ there exist an almost periodic sequence such that each its substring of (suffisiently large) length n has complexity at least αn.

We use a relativized version of the main theorem: we construct $\omega^{(1)}$ satisfying the theorem, then construct $\omega^{(2)}$ satisfying the theorem with oracle $\omega^{(1)}$, then construct $\omega^{(3)}$ satisfying the theorem with oracles $\omega^{(1)}$ and $\omega^{(2)}$, and so on.

After this we apply a special pattern to compose an almost periodic sequence out of all this sequences $\omega^{(1)}, \omega^{(2)}, \ldots$.

Application 2. Generalization of Crieger-Shallit theorem

Fractional powers of strings

$$(01101)^{2.6} = 01101\,01101\,011$$

$$(1001)^{3.75} = 1001\,1001\,1001\,100$$

Fractional power of string is defined only if the denominator of the exponent is equal to the length of the base.

Application 2. Generalization of Crieger-Shallit theorem

Critical exponent

Definition

Let ω be an infinite sequence.

Critical exponent of ω is

$$\sup\{r \mid \exists x \neq \Lambda \colon x^r \text{ is a substring of } \omega\}$$

Cube- and square-free sequences (Thue, Morse and others).

Application 2. Generalization of Crieger-Shallit theorem

Creiger-Shallit theorem:

Any number greater than 1 is a critical exponent of some sequence (in some alphabet than may depend on the number):

$\forall r > 1\, \exists \omega \colon r$ is a critical exponent of ω.

We prove a stronger result:

For any $r > 1$ there exist $\varepsilon > 0$ and a sequence ω such that:

(1) r is a critical exponent of ω and

(2) for any $s > r$ $\quad\omega$ does not contain substring that is $\min\{\varepsilon, \frac{s-r}{s}\}$-close to a power with the exponent s (Hamming distance).

Application 3. Sequences without long periods

Notation. $R(a, l)$ ($a \geq 2$ and $l \geq 1$) is minimal number r such that there exists a sequence in a-letter alphabet without fractional powers x^p for $|x| \geq l$ and $p \geq r$.

Theorem. There exist a number c such that

$$1 + \frac{1}{al} \leq R(a, l) \leq 1 + \frac{c}{al}$$

Here lower bound is a simple combinatorial one (pigeonhole principle) and the upper bound uses the main theorem (not trivial).

Negative results

Generalized Levin Lemma and computable permutations

Lemma (generalized Levin). Assume that $\alpha < 1$. Then there exists an infinite binary sequence ω such that for each numbers k, n

$$\mathrm{K}(\omega_k\omega_{k+1}\ldots\omega_{k+n-1} \mid \omega_0\omega_1\ldots\omega_{k-1}) \geq \alpha n$$

Further generalization is impossible: (For any $\alpha > 0$) there is no sequence such that any its computable permutation satisfies the condition of generalized Levin Lemma (for this $alpha$).

Oracle negative result.

Theorem.

There exist a constant c such that for any two sequences at least one of them contains arbirarily long substrings that has complexity at most c with respect to their position with oracle of the other sequence.

Open question: is there any pair of sequences such that each of them satisfies the condition of Levin Lemma with oracle of the other sequence.

The end