

# Forbidden substrings, Kolmogorov complexity and almost periodic sequences

Andrey Rumyantsev, Maxim Ushakov  
Moscow State University

# Plan

Sequence without  
simple substrings  
(Levin)

- Combinatorial proof
- Multidimensional generalization



Almost periodic sequence without simple prefixes (Muchnik, Semenov, Ushakov)

- Construction of an almost periodic sequence
- Simplified proof



Almost periodic sequence without simple substrings  
(main result)



Multidimensional case

# Forbidden strings

## Combinatorial proposition

### Proposition 1.

Let  $\alpha < 1$ .

For each  $n$  a set  $F_n$  of binary strings of length  $n$  is fixed.

Suppose that  $|F_n| \leq 2^{\alpha n}$ .

Then

$\exists$  a constant  $N$  and an infinite binary sequence  $\omega$

$\forall n > N$

$\omega$  does not have a substring  $x$  that belongs to  $F_n$ .

## Complexity proposition

### Proposition 2. (Levin)

Let  $\alpha < 1$ .

Then

$\exists$  a constant  $N$  and an infinite binary sequence  $\omega$

such that any its substring  $x$  of length greater than  $N$

has high complexity:

$$K(x) > \alpha|x|$$

These two propositions are equivalent.

## Combinatorial proof

**Proposition 3.** (Laslo Lovasz local lemma)

Let  $G$  be a graph with vertex set  $V = \{v_1, \dots, v_n\}$ .

Let  $A_i$  be some event associated with vertex  $v_i$ .

Assume that  $\forall i$  the event  $A_i$  is independent with the random variable “outcomes of all  $A_j$  such that  $v_j$  is not connected to  $v_i$  by an edge”.

Let  $p_i \in (0, 1)$  be a number associated with  $A_i$  in such a way that

$$\Pr[A_i] \leq p_i \prod_{v_j \sim v_i} (1 - p_j)$$

( $v_j \sim v_i$  means that  $v_j$  and  $v_i$  are connected by an edge).

Then

$$\Pr[\text{neither of } A_i \text{ happens}] \geq \prod_{i=1}^n (1 - p_i)$$

and, therefore, this event is non-empty.

We will use this lemma to proof Proposition 1.

## Combinatorial proof (continue)

Assume all bits of  $\omega$  are independent and unbiased.

Vertexes = intervals of length greater than  $N$ .

Edges = pairs of intervals that are not disjoint.

$\forall$  interval  $v$  consider the event  $A_v$ :

“substring of the random string located at  $v$  is forbidden”.

Assume  $p_v = 2^{-\delta|v|}$ .

To apply the lemma, we need to prove that

$$\Pr[A_v] \leq p_v \prod_{\substack{v \text{ and } w \text{ are} \\ \text{not disjoint}}} (1 - p_w).$$

This is true for enough large  $N$ .

## Combinatorial proof (multidimensional)

$d$ -dimensional binary sequence is a map  $\omega : \mathbb{Z}^d \rightarrow \{0, 1\}$ .

### Proposition 4.

Let  $\alpha < 1$ .

For each  $n$  a set  $F_n$  of forbidden cubes with side  $n$  is fixed.

Suppose that  $|F_n| \leq 2^{\alpha n^d}$ .

Then

$\exists$  a constant  $N$  and a  $d$ -dimensional binary sequence  $\omega$

$\forall n > N$

$\omega$  does not have a subcube  $x$  that belongs to  $F_n$ .

This proposition can be proved by the same way as Proposition 1 (using Laslo Lovasz local lemma).

# Almost periodic sequences

## Definition.

$v_0v_1v_2\dots$  is almost periodic if any its substring appears infinitely many times at limited distances,

i.e.,

$\forall$  substring  $x \exists k \forall$  substring  $y$  of length  $k$  contains  $x$ .

## Proposition 5. (Muchnik, Semenov, Ushakov)

Let  $\alpha < 1$ .

Then

$\exists$  an almost periodic sequence  $v_0v_1v_2\dots$

$\forall$  sufficiently large  $k$

$$K(v_0v_1\dots v_{k-1}) > \alpha k$$

Now we combine the results of Levin and Muchnik, Semenov, Ushakov.

## Proposition 6. (main result)

Let  $\alpha < 1$ .

Then

$\exists$  an almost periodic sequence  $v_0v_1v_2\dots$

$\forall$  sufficiently large  $k$

$$\forall n \quad K(v_nv_{n+1}\dots v_{n+k-1}) > \alpha k$$



## Proof of the main result

Almost periodic sequence without simple substrings

Use our construction of almost periodic sequence,  
and  
take bits from Levin-typed sequence instead of random  
bits: we need a sequence from the following proposition.

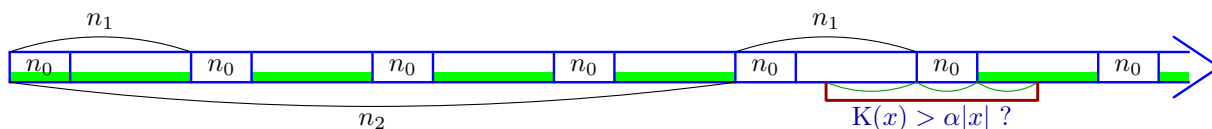
### Proposition 8.

Let  $\alpha < 1$ .

Then  
 $\exists$  an infinite binary sequence  $\omega$

$\forall$  sufficiently large  $k$

$$\forall n \quad K(\omega_n \omega_{n+1} \dots \omega_{n+k-1} \mid \omega_0 \omega_1 \dots \omega_{n-1}) > \alpha k$$



A substring  $x$  of constructed sequence can be split into  
several parts (substrings of  $\omega$ ).

We need the following Lemma to estimate  $K(x)$ .

### Lemma.

If sequence  $\omega$  satisfies the statement of Proposition 8 then

$$K(V(a_0, b_0), V(a_1, b_1), \dots, V(a_{s-1}, b_{s-1})) \geq \\ \alpha L - O(s \log L) - K(a_0 \mid a_1) - \dots - K(a_{s-2} \mid a_{s-1})$$

for any  $a_0 < b_0 \leq \dots \leq a_{s-1} < b_{s-1}$ ,

where  $V(a, b) = \omega_a \omega_{a+1} \dots \omega_{b-1}$

and  $L = (b_0 - a_0) + \dots + (b_{s-1} - a_{s-1})$ .

## Multidimensional case

### Definition.

$d$ -dimensional sequence  $v$  is almost periodic if  
 $\forall$  subcube  $x \exists k \forall$  subcube  $y$  with side  $k$  contains  $x$ .

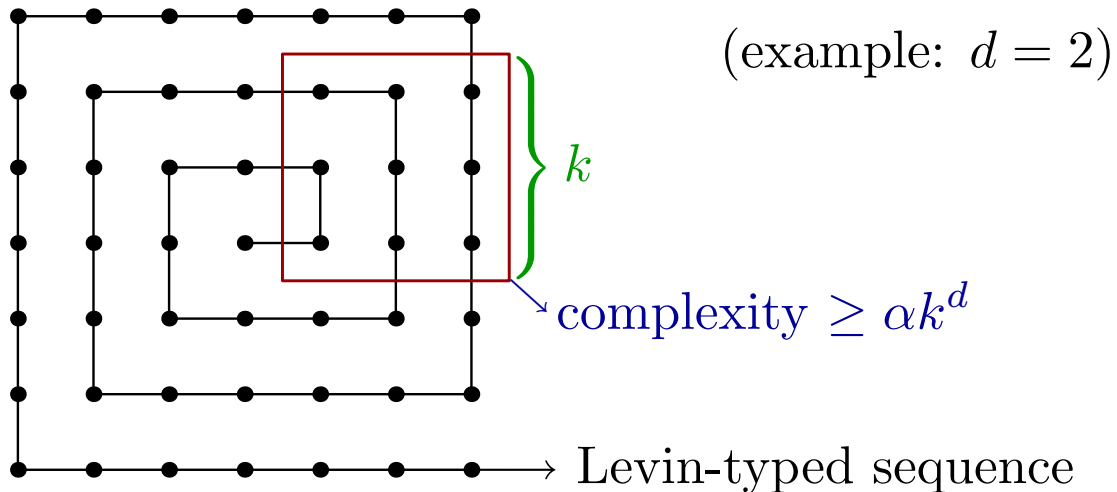
### Proposition 9.

Let  $\alpha < 1$ .

Then

- $\exists$  an almost periodic  $d$ -dimensional sequence  $v$
- $\forall$  sufficiently large  $k$
- $\forall$  subcube with side  $k$  has complexity greater than  $\alpha k^d$ .

Take a sequence from Proposition 8 and write down its terms *along a spiral*:



To get the *almost periodic* sequence:  
*duplicate cubes* by the same way as in 1-dimensional case.

**THE END**