Diophantine flavor of Kolmogorov complexity

Yuri Matiyasevich

St. Petersburg Department of Steklov Institute of Mathematics of Russian Academy of Sciencies

URLS: http://logic.pdmi.ras.ru/~yumat http://www.informatik.uni-stuttgart.de/ifi/ti/personen/Matiyasevich Plan of the talk:

- I. Some definitions and results in the theory of Kolmogorov complexity
- II. Main results in Diophantine computations
- *III. Diophantine flavor of Kolmogorov complexity*

Part I. Some definitions in the theory of Kolmogorov complexity

Part I. Some definitions in the theory of Kolmogorov complexity

360 alternating "1" and "0"

360 alternating "1" and "0"

10101010101010101010101010	110010010000111111011010
101010101010101010101010	101000100010000101101000
101010101010101010101010	110000100011010011000100
101010101010101010101010	110001100110001010001011
101010101010101010101010	10000001101110000011100
101010101010101010101010	110100010010100100000010
101010101010101010101010	010011100000100010001010
101010101010101010101010	011001111100110001110100
101010101010101010101010	0000010000101110111110
101010101010101010101010	101001100011101100010011
101010101010101010101010	100110110010001001010001
101010101010101010101010	010010100000100001111001
101010101010101010101010	100011100011010000000100
101010101010101010101010	110111011110111110010101
101010101010101010101010	000110011011001111001101
360 alternating "1" and "0"	The first 360 binary digits of
	the number π
	-

Without lost of generality, we fix the two-letter alphabet $B = \{0, 1\}$ for both descriptions and words described by them.

Without lost of generality, we fix the two-letter alphabet $B = \{0, 1\}$ for both descriptions and words described by them.

A description mode is a binary relation D between words, i.e., $D \subseteq B^* \times B^*$.

A word A is called a *description* of a word W (according to given description mode D) if $\langle A, W \rangle \in D$.

Without lost of generality, we fix the two-letter alphabet $B = \{0, 1\}$ for both descriptions and words described by them.

A description mode is a binary relation D between words, i.e., $D \subseteq B^* \times B^*$.

A word A is called a *description* of a word W (according to given description mode D) if $\langle A, W \rangle \in D$.

Other terminology/point of view:

D is an *archiving method* A is an *archive* of W D is a compression methodA is a compressed form of W

• $\langle A, W' \rangle \in D \& \langle A, W'' \rangle \in D \Rightarrow W' = W''$ (the uniqueness of the decompression)

- $\langle A, W' \rangle \in D \& \langle A, W'' \rangle \in D \Rightarrow W' = W''$ (the uniqueness of the decompression)
- *D* is recursively enumerable (the effectiveness of decompression)

- $\langle A, W' \rangle \in D \& \langle A, W'' \rangle \in D \Rightarrow W' = W''$ (the uniqueness of the decompression)
- *D* is recursively enumerable (the effectiveness of decompression)
- Every word has at least one description (the universality of archiving method).

- $\langle A, W' \rangle \in D \& \langle A, W'' \rangle \in D \Rightarrow W' = W''$ (the uniqueness of the decompression)
- *D* is recursively enumerable (the effectiveness of decompression)
- Every word has at least one description (the universality of archiving method).

An equivalent point of view: a description mode is an effectively computable function D from some subset \mathcal{P} of B^* onto B^* .

- $\langle A, W' \rangle \in D \& \langle A, W'' \rangle \in D \Rightarrow W' = W''$ (the uniqueness of the decompression)
- *D* is recursively enumerable (the effectiveness of decompression)
- Every word has at least one description (the universality of archiving method).

An equivalent point of view: a description mode is an effectively computable function D from some subset \mathcal{P} of B^* onto B^* .

An possible additional condition of *prefix-free* description mode can be imposed:

• The set $\mathcal{P} = \{A : \exists W \langle A, W \rangle \in D\}$ is prefix-free, i.e., no word from \mathcal{P} is a prefix of another word from \mathcal{P}

Equivalently: if $\langle A, W \rangle \in D$ then $\langle AA', W' \rangle \notin D$ for every word W' and every non-empty word A'.

$$K_D(W) = \min\{|A| : \langle A, W \rangle \in D\}$$

where |A| is the length of A.

$$K_D(W) = \min\{|A| : \langle A, W \rangle \in D\}$$

where |A| is the length of A.

An equivalent definition. The complexity $K_D(W)$ of a word W with respect to a description mode D is defined by

 $K_D(W) = \min\{|A| : D(A) = W\}$

$$K_D(W) = \min\{|A| : \langle A, W \rangle \in D\}$$

where |A| is the length of A.

An equivalent definition. The complexity $K_D(W)$ of a word W with respect to a description mode D is defined by

$$K_D(W) = \min\{|A| : D(A) = W\}$$

A (prefix-free) description mode D is *optimal* if for every (prefix-free) description mode D' there is a number c such that for every word W

 $K_D(W) \le K_{D'}(W) + c.$

$$K_D(W) = \min\{|A| : \langle A, W \rangle \in D\}$$

where |A| is the length of A.

An equivalent definition. The complexity $K_D(W)$ of a word W with respect to a description mode D is defined by

$$K_D(W) = \min\{|A| : D(A) = W\}$$

A (prefix-free) description mode D is *optimal* if for every (prefix-free) description mode D' there is a number c such that for every word W

$$K_D(W) \le K_{D'}(W) + c.$$

Does an optimal (prefix-free) description mode exists?

$$K_D(W) = \min\{|A| : \langle A, W \rangle \in D\}$$

where |A| is the length of A.

An equivalent definition. The *complexity* $K_D(W)$ of a word W with respect to a description mode D is defined by

$$K_D(W) = \min\{|A| : D(A) = W\}$$

A (prefix-free) description mode D is *optimal* if for every (prefix-free) description mode D' there is a number c such that for every word W

$$K_D(W) \le K_{D'}(W) + c.$$

Does an optimal (prefix-free) description mode exist?

Theorem (Kolmogorov-Solomonoff). There are optimal description modes.

Theorem (Kolmogorov-Solomonoff). There are optimal description modes.

Proof. Enumerate all description modes

 $D_0, D_1, \ldots, D_m, \ldots$

viewed as functions from B^* into B^* and construct a *universal set* U:

 $\langle m, A, W \rangle \in U \Leftrightarrow \langle A, W \rangle \in D_m.$

Then define

 $D_{\mathsf{opt}} = \{ \langle \mathbf{1}^m \mathbf{0} A, W \rangle : \langle m, A, W \rangle \in U \}.$

Theorem (Kolmogorov-Solomonoff). There are optimal description modes.

Proof. Enumerate all description modes

 $D_0, D_1, \ldots, D_m, \ldots$

viewed as functions from B^* into B^* and construct a *universal set* U:

 $\langle m, A, W \rangle \in U \Leftrightarrow \langle A, W \rangle \in D_m.$

Then define

$$D_{\mathsf{opt}} = \{ \langle \mathbf{1}^m \mathbf{0} A, W \rangle : \langle m, A, W \rangle \in U \}.$$

Optimal prefix-free description modes also exit.

Theorem (Kolmogorov-Solomonoff). There are optimal description modes.

Proof. Enumerate all description modes

 $D_0, D_1, \ldots, D_m, \ldots$

viewed as functions from B^* into B^* and construct a *universal set* U:

 $\langle m, A, W \rangle \in U \Leftrightarrow \langle A, W \rangle \in D_m.$

Then define

$$D_{\mathsf{opt}} = \{ \langle 1^m 0A, W \rangle : \langle m, A, W \rangle \in U \}.$$

Optimal prefix-free description modes also exit.

We fix a particular optimal (prefix-free) description mode D_{opt} and define Kolmogorov (prefix) entropy of a word W as the complexity of the word W with respect to D_{opt} .

If an infinite word is decidable (i.e., the set of positions of "1" is a decidable set), then the entropy of its initial fragment of length l grows essentially as log(l).

10101010101010101010101010	110010010000111111011010
101010101010101010101010	101000100010000101101000
101010101010101010101010	110000100011010011000100
101010101010101010101010	110001100110001010001011
101010101010101010101010	10000001101110000011100
101010101010101010101010	110100010010100100000010
101010101010101010101010	010011100000100010001010
101010101010101010101010	011001111100110001110100
101010101010101010101010	0000010000101110111110
101010101010101010101010	101001100011101100010011
101010101010101010101010	100110110010001001010001
101010101010101010101010	010010100000100001111001
101010101010101010101010	100011100011010000000100
101010101010101010101010	110111011110111110010101
101010101010101010101010	000110011011001111001101
360 alternating "1" and "0"	The first 360 binary digits of
	the number π
	-

If an infinite word is recursively enumerable (i.e., the set of positions of "1" is recursively enumerable), then the entropy of its initial fragment of length l grows also essentially as log(l).

If an infinite word is recursively enumerable (i.e., the set of positions of "1" is recursively enumerable), then the entropy of its initial fragment of length l grows essentially as log(l).

The entropy of any word cannot be essentially greater than its length. If the entropy of initial fragments of length l of an infinite word W grows as l (up to an additive constant), then this word W can be considered as random sequence of bits.

If an infinite word is recursively enumerable (i.e., the set of positions of "1" is recursively enumerable), then the entropy of its initial fragment of length l grows essentially as log(l).

The entropy of any word cannot be essentially greater than its length. If the entropy of initial fragments of length l of an infinite word W grows as l (up to an additive constant), then this word W can be considered as random sequence of bits.

Gregory Chaitin introduced a particular method of constructing an infinite word of so big prefix-free complexity. This word can be defined as the sequence of binary digits of some real number Ω . This number can be interpreted as the probability that a randomly selected Turing machine stops. Of course, digits of this word do not form a recursively enumerable set. However, there is an effectively computable increasing sequence $\Omega_1, \Omega_2, \ldots, \ldots$ of rational numbers which converges to Ω .

Part II. Main results in Diophantine computations

A Diophantine equation is an equation of the form

 $P(x_1,\ldots,x_m)=0,$

where P is a polynomial with integer coefficients and the unknowns x_1, \ldots, x_m are supposed to be natural numbers.

Part II. Main results in Diophantine computations

A Diophantine equation is an equation of the form

 $P(x_1,\ldots,x_m)=0,$

where P is a polynomial with integer coefficients and the unknowns x_1, \ldots, x_m are supposed to be natural numbers.

A *family* of Diophantine equations is an equation of the form

$$P(a_1,\ldots,a_n,x_1,\ldots,x_m)=0,$$

in which the variables are split into two groups:

- the parameters a_1, \ldots, a_n
- the unknowns x_1, \ldots, x_m

Part II. Main results in Diophantine computations

A Diophantine equation is an equation of the form

$$P(x_1,\ldots,x_m)=0,$$

where P is a polynomial with integer coefficients and the unknowns x_1, \ldots, x_m are supposed to be natural numbers.

A family of Diophantine equations is an equation of the form

$$P(a_1,\ldots,a_n,x_1,\ldots,x_m)=0,$$

in which the variables are split into two groups:

- the parameters a_1, \ldots, a_n
- the unknowns x_1, \ldots, x_m

Consider the set ${\mathfrak M}$ such that

$$\langle a_1,\ldots,a_n\rangle \in \mathfrak{M} \iff \exists x_1\ldots x_m \{P(a_1,\ldots,a_n,x_1,\ldots,x_m)=0\}.$$

Sets which can be defined in this way are called *Diophantine*, the above equivalence being called *Diophantine representation* of the set \mathfrak{M} .

Trivial fact. Every Diophantine set is recursively enumerable.

Trivial fact. Every Diophantine set is recursively enumerable.

Main result (DPRM-theorem). Every recursively enumerable set is Diophantine.

DPRM: after Davis-Putnam-Robinson-Matiyasevich

Non-Deterministic Diophantine Machine (NDDM for short) was introduced by Leonard Adleman and Kenneth Manders.



DPRM-theorem says that NDDMs are as powerful as, say, nondeterministic Turing machines. Two-way bridge







To represent words by natural numbers, we can use two-letter alphabet $\{1,2\}$ and consider words in this alphabet as numbers written in base 2 notation.



$$\langle a_1, \ldots, a_n \rangle \in \mathfrak{M} \iff \exists x_1 \ldots x_m \{ E(a_1, \ldots, a_n, x_1, \ldots, x_m) = 0 \}$$

where E is constructed with addition, subtraction, multiplication and exponentiation.



Every r.e. set has a single-fold exponential Diophantine representation, i.e., such that the x's are uniquely determined by the a's

DPR-theorem. Every r.e. set has an *exponential Diophantine representation*

$$\langle a_1, \ldots, a_n \rangle \in \mathfrak{M} \iff \exists x_1 \ldots x_m \{ E(a_1, \ldots, a_n, x_1, \ldots, x_m) = 0 \}$$

where E is constructed with addition, subtraction, multiplication and exponentiation.



Hilary Putnam's trick: Given a Diophantine equation

$$P(a, b, x_1, \ldots, x_m) = 0$$

we can construct another Diophantine equation of the form

$$Q(a, y_1, \ldots, y_n) = b$$

such that the former equation has a solution in x_1, \ldots, x_m for the same values of the parameters a and b for which the latter equation has a solutions in y_1, \ldots, y_n .

This trick works provided that the parameter b assumes only nonnegative integer values. Part III. Diophantine flavor of Kolmogorov complexity

A description mode is a recursively enumerable relation D between words such that

$$\langle A, W' \rangle \in D \& \langle A, W'' \rangle \in D$$

 $\Rightarrow W' = W''$

Part III. Diophantine flavor of Kolmogorov complexity

A description mode is a recursively enumerable relation *D* between words such that

$$\langle A, W' \rangle \in D \& \langle A, W'' \rangle \in D$$

 $\Rightarrow W' = W''$

description mode is a A Diophantine description mode is a polynoursively enumerable rela- mial P such that

$$P(a, w', x'_1, \dots, x'_k) = P(a, w'', x''_1, \dots, x''_k) = 0$$

\$\Rightarrow w' = w''.

Part III. Diophantine flavor of Kolmogorov complexity

A description mode is a recursively enumerable relation D between words such that

$$\langle A, W' \rangle \in D \& \langle A, W'' \rangle \in D$$

 $\Rightarrow W' = W''$

A Diophantine description mode is a polyno-
mial P such that
$$P(a, w', x'_1, \dots, x'_k) = P(a, w'', x''_1, \dots, x''_k) = 0$$
$$\Rightarrow w' = w''.$$

The complexity $K_D(W)$ of a word W is the length of the shortest compressed form of W:

 $K_D(W) = \min\{|A| : \langle A, W \rangle \in D\}.$

The Diophantine complexity $D_P(w)$ of a number w with respect to such a P is defined by

$$D_P(w) = \min\{|a| : P(a, w, x_1, \dots, x_k) = 0\}.$$

A description mode D is an effectively computable function D from some subset \mathcal{P} of B^* onto B^*

fectively computable function Dfrom some subset \mathcal{P} of B^* onto B^*

A description mode D is an ef- \parallel A Diophantine description mode is a polynomial P which assumes every natural number as its value and such that

$$P(a, x'_1, \dots, x'_k) \ge 0 \& P(a, x''_1, \dots, x''_k) \ge 0$$

$$\Rightarrow P(a, x'_1, \dots, x'_k) = P(a, x''_1, \dots, x''_k).$$

fectively computable function Dfrom some subset \mathcal{P} of B^* onto B^*

A description mode D is an ef- \parallel A Diophantine description mode is a polynomial P which assumes every natural number as its value and such that

$$P(a, x'_1, \dots, x'_k) \ge 0 \& P(a, x''_1, \dots, x''_k) \ge 0$$

$$\Rightarrow P(a, x'_1, \dots, x'_k) = P(a, x''_1, \dots, x''_k).$$

shortest compressed form of W:

$$K_D(W) = \min\{|A| : D(A) = W\}.$$

The complexity $K_D(W)$ of a || The Diophantine complexity $D_P(w)$ of a word W is the length of the number w with respect to such a P is defined bv

$$D_P(w) = \min\{|a| : P(a, x_1, \dots, x_k) = w\}.$$

A Diophantine description mode is a polynomial P such that for every w the inequality

$$P(w', x_1', \ldots, x_k') \ge 0$$

has a solution and $P(w', x_1', \dots, x_k') = P(w'', x_1'', \dots, x_k'') \ge 0 \Rightarrow w' = w''.$

The Diophantine complexity $D_P(w)$ of a number w with respect to such a P is defined by

 $D_P(w) = \min\{|P(w, x_1, \dots, x_k)| : P(w, x_1, \dots, x_k) \ge 0\}.$

 $K_D(W) \le K_{D'}(W) + c.$

A Diophantine description mode P is optimal if for every Diophantine description mode P' there is a number csuch that

 $K_P(w) \le K_{P'}(w) + c.$

 $K_D(W) \le K_{D'}(W) + c.$

A Diophantine description mode P is optimal if for every Diophantine description mode P' there is a number csuch that

 $K_P(w) \le K_{P'}(w) + c.$

Theorem. There are optimal Diophantine description modes.

 $K_D(W) \le K_{D'}(W) + c.$

A Diophantine description mode P is optimal if for every Diophantine description mode P' there is a number csuch that

 $K_P(w) \le K_{P'}(w) + c.$

Theorem. There are optimal Diophantine description modes.

Proof 1. Pass the bridge from left to right.

 $K_D(W) \le K_{D'}(W) + c.$

A Diophantine description mode P is optimal if for every Diophantine description mode P' there is a number csuch that

 $K_P(w) \le K_{P'}(w) + c.$

Theorem. There are optimal Diophantine description modes.

Proof 1. Pass the bridge from left to right.

Proof 2. Pass the bridge from right to left—use the existence of a universal Diophantine equation:

$$\exists y_1 \dots y_n U(a, k, y_1, \dots, y_n) = 0 \Leftrightarrow \exists x_1 \dots x_{m_k} P_k(a, x_1, \dots, x_{m_k}) = 0$$

A description mode D is optimal if \parallel for every description mode D' there $\parallel optimal$ if for every Diophantine deis a number c such that for every word W

 $K_D(W) \le K_{D'}(W) + c.$

A Diophantine description mode P is scription mode P' there is a number csuch that

 $K_P(w) \le K_{P'}(w) + c.$

Theorem. There are optimal Diophantine description modes.

Theorem (Matiyasevich). Every optimal Diophantine description mode can be defined as

$$a = \min\{P_{\mathsf{opt}}(w, x_1, \dots, x_m)\}$$

where P_{opt} is a suitable polynomial assuming only non-negative values.

Chaitin's unpredictable Ω can be interpreted as the probability that a randomly selected Turing machine stops interpreted as the probability that a randomly selected Turing machine stops

Chaitin's unpredictable Ω can be Chaitin's unpredictable Ω can be interpreted as the probability that a randomly selected Diophantine equation has a solution

interpreted as the probability that a randomly selected Turing machine stops

Chaitin's unpredictable Ω can be Chaitin's unpredictable Ω can be interpreted as the probability that a randomly selected Diophantine equation has a solution

There is an effectively computable increasing sequence $\Omega_1, \Omega_2, \ldots, \ldots$ of rational numbers which converges to Ω .

chine stops

Chaitin's unpredictable Ω can be Chaitin's unpredictable Ω can be interinterpreted as the probability that preted as the probability that a randomly a randomly selected Turing ma- || selected Diophantine equation has a solution

verges to Ω .

There is an effectively computable || Chaitin [1975] constructed particuincreasing sequence $\Omega_1, \Omega_2, \ldots, \ldots$ | lar exponential Diophantine equation of rational numbers which con- $|| E(k, z_1, \dots, z_m) = 0$ such that the k-th digit of Ω is equal to 1 iff the equation has infinitely many solutions in z_1, \ldots, z_m .

a randomly selected Turing machine stops

verges to Ω .

There is an effectively computable increasing sequence $\Omega_1, \Omega_2, \ldots, \ldots$ of rational numbers which converges to Ω .

Chaitin's unpredictable Ω can be || Chaitin's unpredictable Ω can be interinterpreted as the probability that \parallel preted as the probability that a randomly selected Diophantine equation has a solution

There is an effectively computable Chaitin [1975] constructed particuincreasing sequence $\Omega_1, \Omega_2, \ldots, \ldots$ | lar exponential Diophantine equation of rational numbers which con- $|| E(k, z_1, \dots, z_m) = 0$ such that the k-th digit of $\boldsymbol{\Omega}$ is equal to 1 iff the equation has infinitely many solutions in z_1, \ldots, z_m .

> Toby Ord and Tien D. Kieu [2003] constructed particular exponential Diophantine equation $E(k, z_1, \ldots, z_m) = 0$ such that for every k this equation has finitely many solutions in z_1, \ldots, z_m and the number of solutions is odd iff the k-th digit of Ω is equal to 1.

Theorem (Matiyasevich [2003]). Let \mathfrak{C} be any infinite decidable set with infinite complement. Then we can construct an exponential Diophantine equation $E(k, z_1, \ldots, z_m) = 0$ such that

- for every k the equation has finitely many solutions in z_1, \ldots, z_m ;
- the number of solutions of the equation for givel value of k belongs to the set C iff the k-th digit of Ω is equal to 1.

Theorem (Matiyasevich [2003]). Let \mathfrak{C} be any infinite decidable set with infinite complement. Then we can construct an exponential Diophantine equation $E(k, z_1, \ldots, z_m) = 0$ such that

- for every k the equation has finitely many solutions in z_1, \ldots, z_m ;
- the number of solutions of the equation for givel value of k belongs to the set \mathfrak{C} iff the k-th digit of Ω is equal to 1.

Open Problem. Could we improve any of the above mentioned definitions of Ω via the number of solutions of exponential Diophantine equations to similar definitions via the number of solutions of genuine (i.e., without exponentiation) Diophantine equations?