Kolmogorov Complexity

TP 2024/25

Szymon Wojtulewicz

December 5, 2025

Overview

1. **Motivation**
2. **Description Methods**
3. **Formal definition**
4. **Incompressibility**

Image compression

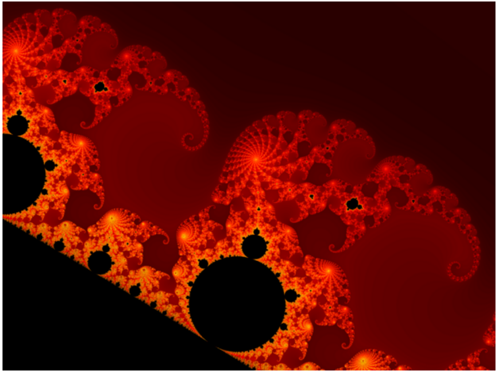


Figure: https://commons.wikimedia.org/w/index.php?curid=13441999

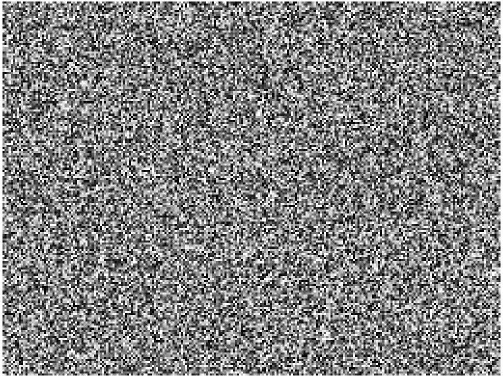
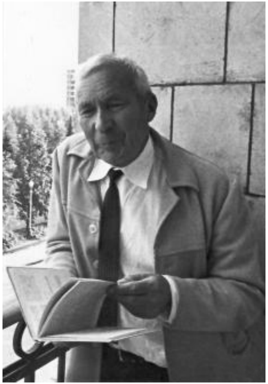


Figure: https://commons.wikimedia.org/w/index.php?curid=24614072

String compression

$$A = 4444444444$$
$$B = 2718281828$$
$$C = 1756475382$$

Andrej Kolmogorov



Андре́й Никола́евич Колмого́ров

1903-1987

1963

Figure:
https://commons.wikimedia.org/w/index.php?curid=11829175

Complexity of description methods

Description Method

$$f : \Sigma^* \to \Sigma^* \quad \Sigma = \{0, 1\}$$

Complexity of string x under description method f

$$C_f(x) := \min\{|p| : f(p) = x\}$$

Minorization

Minorization

A partial function $f : \{0,1\}^* \to \{0,1\}^*$ minimizes a partial function $g : \{0,1\}^* \to \{0,1\}^*$ if there exists a constant c such that for all strings x,

$$C_f(x) \leq C_g(x) + c.$$

Remark

To encode the pair $\langle x, y \rangle$, use $2|x| + |y| + 2$ bits: repeat each bit of x twice, add the string 01, and then y.

Universal Description Method

Theorem: There exists a partial recursive function U that minorizes every partial recursive function

For a universal TM U we have:

$$C_U(x) = \min\{|\langle M, w\rangle| : M \text{ halts on input } w \text{ and outputs } x\}.$$

Given any partial recursive function g, let M be a TM that computes it. Let x be any string, and let p be a shortest string such that $M(p) = x$. $U(\langle M, p\rangle) = x$.

$$|\langle M, p\rangle| = c_M + |p|,$$

where $c_M = 2|\langle M\rangle| + 2$. c_M is independent of p.

$$C_U(x) \leq |\langle M, p\rangle| \leq c_M + |p| = c_M + C_g(x).$$

$\square$

Kolmogorov Complexity

Kolmogorov Complexity

The Kolmogorov complexity of a string x is

$$C(x) := C_U(x) = \min\{|\langle M, w\rangle| : \text{TM } M \text{ halts on input } w \text{ and outputs } x\}.$$

Incompressibility

Incompressible Strings

A string x is *incompressible* or *Kolmogorov random* if $C(x) \geq |x|$.

Remark

There exist incompressible strings of every length: there are 2^n binary strings of length n, but only $2^n - 1$ binary strings of length strictly less than n.

Undecidability of incompressibility

It is undecidable whether a string is incompressible

On large input n take the first incompressible string of such length and make this short sentence its description. A contradiction □

Exercise

Prove the above using a reduction from the Halting Problem.

Prime numbers

Exercise

There are infinitely many prime numbers

The End