L17 2022-10-28

Randomness.

How do we tell if a string is random?

e.g. 01010101010101010101010101
vs 01000111010100111110010100

The first can be compressed as "write 01 12 times".

A more fair representation is "$\langle M, y\rangle$" st $M(y) = x$.

def Kolmogorov cplex of x $K(x) = \min_{M(y)=x} |\langle M, y\rangle|$.

K does not increase string size:

th: $\exists c\ \forall x\ K(x) \leq |x| + c$.

proof: Let M be the TM that does nothing.

K is optimal:

th let $F: \{0,1\}^* \to \{0,1\}^*$ be a computable encoding, and

let $K_F(x) = \min_{F(y)=x} |y|$. Then $\exists c$ st $K(x) \leq K_F(x) + c$.

proof: Let M be the TM that computes F.

A string is random if it is incompressible. $(K(x) \geq x)$

th Incompressible strings exist.

proof: There are 2^n strings of length n.
But only $1 + 2 + \dots + 2^{n-1} = 2^n - 1$ representations
of length $< n$.

th 99.9% of strings cannot be compressed more than 10 bits.

th If $\langle M, y\rangle$ is a witness for $K(x)$, then $\langle M, y\rangle$ is incompressible.

proof: let $\langle N, z\rangle$ be a witness for $\langle M, y\rangle$.

Let P be the TM:
simulate input tape ← produces $\langle M, y\rangle$
simulate input tape ← produces $\langle x\rangle$

$\langle P, \langle N, z\rangle\rangle$ shorter witness for $K(x)$.

th K is not computable.

proof: sup M is a TM computing K.

then we build a TM that compresses large strings!

Let N be the TM:
for each string s:
if $K(s) > M + 2000$
return s.

$|N| \leq |M| + 1000$, hence $\langle N, ``\rangle$ is a witness that

$K(s) \leq |M| + 1100$. Contradiction!!