TI-Lecture-24-3-2020

24 March, 2020 13:54

3333333 ... 3 ← $\lg n$

3 1 4 1 5 9 2 6 5 2 5 ← π $\lg n$

8 4 3 7 7 2 9 3 1 5 2 ← random? n

Kolmogorov complexity

- Solomonoff $\approx$ '56
- Chaitin

Why is something non-random.

- predictable $\approx$
- easy to describe

string $\longrightarrow$ complexity

$\{0,1\}^* \to \mathbb{N}$

string $\longrightarrow$ "length of its description"

$f : \{0,1\}^* \to \{0,1\}^*$ ... language

description string

f ... is partially computable
... $\exists$ an algorithm for f.
which might not halt
on some descriptions

$$f(d) = \underline{\underline{x}}$$

Def: Kolmogorov complexity of string $x \in \{0,1\}^*$
relative to f is

$$C_f(x) = \min_{d \in \{0,1\}^*} \{ |d| \; ; \; f(d) = x \}$$

Thm: $\exists$ universal partially computable $\mathcal{U} : \{0,1\}^* \to \{0,1\}^*$
$\forall$ p.r.f f $\exists c_f \in \mathbb{N}$ $\forall x \in \{0,1\}^*$

$$C_{\mathcal{U}}(x) \leq C_f(x) + c_f \, .$$

Pf:

$$\phi_1, \phi_2, \phi_3, \ldots$$

$\uparrow$
description of a TM

program for f
$\to$ its index

pairing fcn $\langle , \rangle : \{0,1\}^* \times \{0,1\}^* \to \{0,1\}^*$

$i, x \in \{0,1\}^*$ $\quad \langle i, x \rangle = \underline{0^{|i|} 1 \, i \, x}$

- $\mathcal{U}$ on input w separates w into i & d s.t. $w = \langle i, d \rangle$

  run $\phi_i(d)$ & output whatever $\phi_i(d)$ outputs.

Pick f

$f = \phi_j$ for some j.

$c_f = 2|j|+1$

$\forall x \in \{0,1\}^*$ $\quad d \quad f(d) = x$

$\mathcal{U}(\langle j, d \rangle) = x$

$\Rightarrow \quad C_{\mathcal{U}}(x) \le 2|j|+1 + C_f(x)$ $\quad \boxed{QED}$

Fix universal p.r.f. $\mathcal{U}$

and we define $C(x) = C_{\mathcal{U}}(x)$.

Ex: 1) $C(x) \le |x| + O(1)$

$f(x) = x$ ... identity

$C_f(x) \le |x|$

2) $C(0^n) \le \lg n + O(1)$

$C(0^n) \underset{\text{"="}}{\le} C(\overline{n}) + O(1)$

binary encoding of n.

3) $C(\pi_{1..n}) = C(0^n) + O(1)$

4)

$\forall n \, \exists x \in \{0,1\}^n \quad C(x) \ge |x| = n$

$\forall n \; \exists x \in \{0,1\}^n \qquad C(x) \geq |x| \ldots$

$$2^i \qquad \sum_{i=0}^{n-1} 2^i = 2^n - 1 \qquad 2^n = |\{0,1\}^n|$$

$\Rightarrow$ # of descriptions of legth $\leq n-1$.

- cannot find those strigs algorithmically
- cannot compute $C(x)$ by an algoritm.
- many such strings.

$x \qquad C(x) \geq |x| \quad \ldots \; x$ is Kolmogorov random.

$C(x)$ ... "the length of shortest pgm. to print out x".

$\Rightarrow$ measures amount of "algorithmic" information in an individual strig.

$H(X)$

Conditional Kolmogorov complexity

$x, y \in \{0,1\}^* \qquad$ p.r.f. f

$$C_f(x \mid y) = \min \{ |d| ; \; d \in \{0,1\}^*, \; f(\langle d, y \rangle) = x \}$$

Thm: $\exists \mathcal{U} ; \; \forall$ p.r.f. $f \quad \exists c_f \quad \forall x, y \in \{0,1\}^*$

$$C_u(x|y) \le C_f(x|y) + c_f$$

Pf: the same as above $\boxed{}$

Fix U; $\quad C(x) \equiv C(x|\varepsilon)$ ← empty string

; $C(x,y) \equiv C(\langle x,y \rangle)$

Thm: $\forall x,y \quad C(x,y) \le C(x) + C(y|x) + O(\lg C(x))$

Pf: need to describe $\langle x,y \rangle$

$d_x + d_{y|x}$ + description of length of d_x

$|d_x| = C(x)$

$\boxed{}$

Thm: $\forall x,y \quad C(x,y) \ge C(x) + C(y|x) - O(\lg C(x,y))$.

true, but we will not prove it.

Kolmogorov information (of y about x)

$$I_c(x:y) = C(x) - C(x|y)$$

symmetric: $\quad I_c(x:y) = I_c(y:x) \pm O(\lg(C(x,y)))$

Ex: $\forall n \quad \exists x \in \{0,1\}^n \quad C(x|n) \ge n$

$c(n) \ge \lg n \qquad \quad \Gamma \ldots n \ldots x$

$$I(x:n) = C(x) - C(x|n) = O(1)$$

$$I(n:x) = \underbrace{C(n)}_{\geq \lg n} - \underbrace{C(n|x)}_{O(1)} \geq O(\lg n)$$

X r.v. $\{0,1\}^n$ $\quad H(X)$

$E[C(X)]$

$\underline{H(X) \approx E[C(X)]}$ false

Kolm. random x $\quad \Pr[X = x] = 1$

<u>Thm</u>: r.var's $X_1, X_2, X_3 \ldots$ (computable sequence)

X_n on $\{0,1\}^n$

$\forall n$

$$H(X_n) - O(\lg n) \leq E[C(X_n)] \leq H(X_n) + O(\lg n).$$

Computable sequence $\equiv$ $n, x \in \{0,1\}^n \rightarrow \Pr[X_n = x]$

<u>Pf</u>: $E[\underline{C(X_n)}] \leq H(X_n) + O(\lg n)$.

goal: make $E[C(X_n)] \leq H(X_n) + \ldots$

by designing descriptions for strings in X_n

$X_n \longrightarrow$ $\underline{\text{Shannon code}}$ $\quad d_x$

$\underline{x \in X_n}$ $\quad \ell(x)$ $\quad E[\ell(x)] \leq H(X_n) + 1$

$, \; f$

description for x w.r.t. some p.r.f. f

alg for f:

- $d'_x = \langle n, d_x \rangle$ ... calculate X_n
- build the Shannon code for X_n
- output x corresponding to d_x.

$$\mathbb{E}[C_f(X_n)] \leq \mathbb{E}[O(\lg n) + \ell(X_n)]$$
$$\leq \mathbb{E}[\ell(X_n)] + O(\lg n)$$
$$\leq H(X_n) + O(\lg n). \checkmark$$

$$H(X_n) \leq \mathbb{E}[C(X_n)] + O(\lg n)$$

Pbl. 4 in HW.

$\blacksquare$