Model of Computation

i) read input from left to right
ii) has a working tape
iii) write output from left to right

Universal Computer $\mathcal{U}$

- for each computer $\mathcal{A}$, there is a program $s_{\mathcal{A}}$, satisfying

$$\mathcal{U}: s_{\mathcal{A}} \cdot p \mapsto f_{\mathcal{A}}(p)$$

For a computer $\mathcal{A}$, it def.s a func.

$$f_{\mathcal{A}} : \{0,1\}^* \to \{0,1\}^* \cup \{0,1\}^\infty$$

Def) The Kolmogrov complexity $K_{\mathcal{U}}(x)$ of a string x w.r.t. a uni. comp. $\mathcal{U}$ is

$$K_{\mathcal{U}}(x) := \min_{p : \mathcal{U}(p) = x} l(p)$$

Def) The Conditional ~

$$K_{\mathcal{U}}(x \mid l(x)) := \min_{p : \mathcal{U}(p, l(x)) = x} l(p)$$

(my understand: there is some mapping $(p, l) \mapsto p_l$.)

Thm) (Universality of $K_{\mathcal{U}}$)

$\mathcal{U}$ is univ. $\to \forall \mathcal{A}$: computer. $\exists c_{\mathcal{A}}$. $\forall x$. $K_{\mathcal{U}}(x) \le K_{\mathcal{A}}(x) + c_{\mathcal{A}}$

Corollary) $\mathcal{U}, \mathcal{A}$ are univ. $\to \exists c. \forall x. |K_{\mathcal{U}}(x) - K_{\mathcal{A}}(x)| < c$

pf) $f_{\mathcal{A}}(p_{\mathcal{A}}) = x$, $l(p_{\mathcal{A}}) = K_{\mathcal{A}}(x)$

$l(s_{\mathcal{A}} \cdot p_{\mathcal{A}}) = l(s_{\mathcal{A}}) + l(p_{\mathcal{A}}) = c_{\mathcal{A}} + l(p_{\mathcal{A}})$

$K_{\mathcal{U}}(x) = \min_{p : \mathcal{U}(p) = x} l(p) \le l(s_{\mathcal{A}} \cdot p_{\mathcal{A}}) = c_{\mathcal{A}} + l(p_{\mathcal{A}}) = c_{\mathcal{A}} + K_{\mathcal{A}}(x)$

Thm. $K(x \mid l(x)) \le l(x) + c$

pf. Print(x)

Thm. (Ubd of K)

$$K(x) \le K(x \mid l(x)) + \log^* l(x) + c$$

$\log^* n := \log n + \log\log n + \cdots$

pf. desc. $l(x)$ as " $\cdots$ [$\log\log l(x)$] [$\log l(x)$ bits] "

" Thm. (Lbd of K) $|\{x \in \{0,1\}^* \mid K(x) < k\}| < 2^k$ "

pf) $|\{x \mid K(x) < k\}| \le \{p \mid l(p) < k\} < 2^k$

Notation : $H_0(p) := -p\log p - (1-p)\log(1-p)$

That is, $H_0(X)$ for a rand. var. X, is a rand. var.

Fact :

$$\sqrt{\frac{n}{8k(n-k)}} \cdot 2^{nH(k/n)} \le \binom{n}{k} \le \sqrt{\frac{n}{\pi k(n-k)}} \cdot 2^{nH(k/n)}$$

Example (Seq. of n-bits w/ k ones)

p: Print the i-th of seq.s of n-bits w/ k ones

$$\to l(p) = c + \underbrace{\log n}_{\text{to express } k} + \underbrace{\log \binom{n}{k}}_{\text{to expr. } i} \quad (n \text{ is known})$$

$$\le c' + \log n + nH(k/n) - \tfrac{1}{2}\log n$$

$$= c' + \tfrac{1}{2}\log n + nH(k/n)$$

$$\log\left(\sqrt{\frac{n}{\pi k(n-k)}} \cdot 2^{nH(k/n)}\right)$$

$$= \log\left(\frac{1}{\sqrt{\pi n p q}} \cdot 2^{nH(k/n)}\right) \quad (p = \tfrac{k}{n},\ q = 1-p)$$

$$= -\tfrac{1}{2}\log n - \tfrac{1}{2}\log pq + nH(k/n) + c_0$$

$$\le -\tfrac{1}{2}\log n + nH(k/n) + c_0'$$

" Thm $K(x^{(n)} \mid n) \le n H_0\left(\frac{1}{n}\sum x_i\right) + \frac{1}{2}\log n + c$ "

Lem. $\forall \mathcal{U}$: computer, $\sum_{p:\mathcal{U}(p)\text{ halts}} 2^{-l(p)} \leq 1$ pf) $\{p \mid \mathcal{U}(p) \text{ halts}\}$ is prefix-free

Thm (Rel. of K and H)

$\{X_i\}$, $X_i \sim f(x)$, i.i.d. $x \in \mathcal{X}$, $|\mathcal{X}| < \infty$.

Let $f(x^{(n)}) = \prod f(x_i)$

$$\exists c.\ H(X) \leq \frac{1}{n} \sum_{x^{(n)}} f(x^{(n)}) K(x^{(n)} \mid n) \leq H(X) + \frac{(|\mathcal{X}|-1)\log n}{n} + \frac{c}{n}$$

That is,

"$E(\frac{1}{n} \cdot K(x^{(n)} \mid n)) \to H(X)$"

pf) $(H(X) \leq \frac{1}{n} \sum_{x^{(n)}} f(x^{(n)}) K(x^{(n)} \mid n))$

$$\sum_{x^{(n)}} f(x^{(n)}) K(x^{(n)} \mid n) \underset{\substack{\downarrow \\ c:\, x^{(n)} \mapsto p \in \{0,1\}^*,\ \text{prefix-free}}}{=} \sum_{x^{(n)}} f(x^{(n)}) \cdot |c(x^{(n)})| = E|c(X_1 \cdots X_n)| \underset{\substack{\downarrow \\ \text{Theory of} \\ \text{Source coding}}}{\leq} H(X_1, \cdots, X_n) = nH(X)$$

$$\left(\frac{1}{n} \sum_{x^{(n)}} f(x^{(n)}) K(x^{(n)} \mid n) \leq H(X) + \frac{(|\mathcal{X}|-1)\log n}{n} + \frac{c}{n}\right)$$

Sps X_i is bin.

$$K(x^{(n)} \mid n) \leq nH_0\left(\frac{1}{n}\sum x_i\right) + \frac{1}{2}\log n + c$$

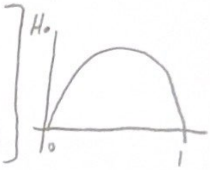
$$\begin{aligned} EK(X_1 \cdots X_n \mid n) &\leq n\, EH_0\left(\frac{1}{n}\sum X_i\right) + \frac{1}{2}\log n + c \\ &\leq nH_0\left(\frac{1}{n}\sum EX_i\right) + \frac{1}{2}\log n + c \\ &= nH_0(\theta) + \frac{1}{2}\log n + c \quad (X_1, \cdots, X_n \text{ are iid.} \sim \text{Bernoulli}(\theta)) \end{aligned}$$

[Jensen's inequ.

$E(\varphi(X)) \leq \varphi(E(X))$ when φ is concave

Concativity of H_0

$H_0(\lambda p_1 + (1-\lambda)p_2) \geq \lambda H_0(p_1) + (1-\lambda) H_0(p_2)$]



If $\mathcal{X}$ is not bin.

$$K(x^{(n)} \mid n) \leq nH(P_{x^{(n)}}) + (|\mathcal{X}|-1)\log n + c$$

(# elts of $P_{x^{(n)}}$-type $\leq 2^{nH_0(P_{x^{(n)}})}$ (#i)-1 # of i-th symbol)

$\downarrow$ $\to$?

$$EK(X^n \mid n) \leq nH(X) + (|\mathcal{X}|-1)\log n + c$$

Corollary) $E\frac{1}{n}K(n^{(n)}) \to H(X)$

Def. For integer n, $K(n) = \min_{p:\, \mathcal{U}(p)=n} l(p)$

Thm. $K_{\mathcal{U}}(n) \leq K_{\mathcal{A}}(n) + c_{\mathcal{A}}$, $\mathcal{U}, \mathcal{A}$: univ.

Thm. $K(n) \leq \log^* n + c$

Thm. There are an inf # of ints n s.t. $K(n) > \log(n)$

pf) Note that $\sum_n 2^{-K(n)} \leq 1$.

Sps $K(n) < \log n$ for all $n > n_0$, then

$$\sum_{n=n_0}^{\infty} 2^{-K(n)} > \sum_{n=n_0}^{\infty} 2^{-\log n} = \sum_{n=n_0}^{\infty} \frac{1}{n} = \infty \ \lightning$$

Def) $x^{(n)}$ is algorithmically random if $K(x^{(n)}|n) \geq n$

Def) $x \in \{0,1\}^\infty$ is incompressible if $\lim \frac{K(x_1 \cdots x_n|n)}{n} = 1$

Thm) $X_1, \cdots, X_n \sim \text{Bernoulli}(\frac{1}{2})$

$$P(K(X_1 X_2 \cdots X_n|n) < n-k) < 2^{-k}$$

pf) $P(K(X_1 \cdots X_n|n) < n-k)$

$$= \sum_{x^{(n)}: K(x^{(n)}) < n-k} 2^{-n}$$

$$= |\{x^{(n)} \mid K(x^{(n)}) < n-k\}| \cdot 2^{-n}$$

$$< 2^{n-k} \cdot 2^{-n} = 2^{-k}$$

Thm) (Strong law of large #s for incompressible seq.)

$x_1 x_2 \cdots$ is incomp. implies $\frac{1}{n}\sum_1^n x_i \to \frac{1}{2}$

pf) Let $\theta_n = \frac{1}{n}\sum^n x_i$

typo in the book: $\frac{2\log n}{n}$

$$\frac{K(x^{(n)}|n)}{n} < H_0(\theta_n) + \frac{\log n}{2n} + \frac{c'}{n}$$

Since $x_1 \cdots$ is incompr.,

$$\forall \varepsilon. \exists n.\ 1 - \varepsilon \leq \frac{K(x^{(n)}|n)}{n} \leq H_0(\theta_n) + \frac{\log n}{2n} + \frac{c'}{n}$$

Thus

$$\forall \varepsilon. \exists n.\ H_0(\theta_n) > 1 - \frac{\frac{1}{2}\log n + c'}{n} - \varepsilon$$

$$\Rightarrow \forall \varepsilon. \exists n.\ H_0(\theta_n) > 1 - \varepsilon$$

$$\Rightarrow \forall \varepsilon. \exists n.\ \theta_n \in (\tfrac{1}{2} - \varepsilon, \tfrac{1}{2} + \varepsilon)$$

Recall) $\{X_i\}$: i.i.d.

$$E\frac{1}{n}K(X^n|n) \to H(X)$$

They do not imply each other.

Thm $\{X_i\} \sim$ i.i.d. Bernoulli(θ)

$\frac{1}{n}K(X^n|n) \to H_0(\theta)$ in probability

pf) Let $\overline{X}_n = \frac{1}{n}\sum X_i$

typo in the book: $2\log n$

$$K(X^n|n) \leq n H_0(\overline{X}_n) + \frac{1}{2}\log n + c$$

By weak law,

$\overline{X}_n \to \theta$ in prob. $(\equiv \lim \Pr(|\overline{X}_n - \theta| < \varepsilon) = 1$ for all $\varepsilon > 0)$

Thus, $\lim \Pr(K(X^n|n) \leq n H_0(\theta) + \frac{1}{2}\log n + c + \varepsilon) = 1\ \forall \varepsilon > 0$

$$\Rightarrow \lim \Pr(\tfrac{1}{n}K(X^n|n) \leq H_0(\theta) + \varepsilon) = 1,\ \forall \varepsilon > 0$$

Now we have to show

$$\equiv \lim \Pr(H_0(\theta) - \tfrac{1}{n}K(X^n|n) \leq \varepsilon) = 1,\ \forall \varepsilon > 0.$$

Let $A_\varepsilon^{(n)}$ be a typical set

$$A_\varepsilon^{(n)} := \{ x^{(n)} \mid |-\tfrac{1}{n}\log p(x^{(n)}) - H(X)| < \varepsilon \}$$

Then

i) $|A_\varepsilon^{(n)}| \geq (1-\varepsilon) 2^{n(H(X)-\varepsilon)}$ ii) $\Pr[x^{(n)} \in A_\varepsilon^{(n)}] \geq 1-\varepsilon$

$\{x^{(n)} | K(x^{(n)}|n) < n(H_0(\theta)-c)\}| \leq 2^{n(H_0(\theta)-c)}$

Lbd of K

For any fixed c, For any $\varepsilon > 0$, $\exists$ large n

$$\begin{aligned}
&\Pr(K(X^{(n)}|n) < n(H_0(\theta)-c)) \\
&\quad \leq \Pr(X^{(n)} \notin A_\varepsilon^{(n)}) + \Pr(X^{(n)} \in A_\varepsilon^{(n)}, K(X^{(n)}|n) < n(H_0(\theta)-c)) \\
&\quad \leq \varepsilon + \sum_{x^{(n)} \in A_\varepsilon^{(n)}, K(x^{(n)}|n) < nH_0(\theta)-c} p(x^{(n)}) \qquad \text{— ii)} \\
&\quad \leq \varepsilon + \sum_{\substack{x^{(n)} \in A_\varepsilon^{(n)}, \\ K(x^{(n)}|n) < nH_0(\theta)-c}} 2^{-n(H_0(\theta)-\varepsilon)} \qquad \text{— def} \\
&\quad \leq \varepsilon + 2^{n(H_0(\theta)-c)} \cdot 2^{-n(H_0(\theta)-\varepsilon)} \qquad \text{— iii)} \\
&\quad = \varepsilon + 2^{-n(c-\varepsilon)}
\end{aligned}$$

That is, $\Pr(K(X^{(n)}|n) < n(H_0(\theta)-c))$

$$= \Pr(c < H_0(\theta) - \tfrac{1}{n}K(X^{(n)}|n)) \to 0 \text{ as } n \to \infty \qquad \square$$

Def) The universal probability of a string x (on an universal comp).

$$P_{\mathcal{U}}(x) = \sum_{p:\mathcal{U}(p)=x} 2^{-l(p)} = \Pr(\mathcal{U}(p)=x).$$

Thm For every computer $\mathcal{A}$,

$$P_{\mathcal{U}}(x) \geq c'_{\mathcal{A}} P_{\mathcal{A}}(x) \qquad c'_{\mathcal{A}} \text{ dep. on } \mathcal{A} \text{ and } \mathcal{U}$$

pf)

$$P_{\mathcal{U}}(x) = \sum_{p:\mathcal{U}(p)=x} 2^{-l(p)} \geq \sum_{p':\mathcal{A}(p')=x} 2^{-l(p')-c_{\mathcal{A}}} = c'_{\mathcal{A}} P_{\mathcal{A}}(x).$$

↳ A의 설계도가 U에 입력된 확률

Def. (Chaitin's #)

$$\Omega := \sum_{\substack{p:\mathcal{U}(p) \\ \text{halts}}} 2^{-l(p)} = \Pr_{p \sim \text{Bernoulli}(\frac{1}{2})} [\mathcal{U}(p) \text{ halts}].$$

대충 망가짐 vs 안멈춤 vs 멈춤점 되돌

i) Ω is noncomputable

ii) Ω is a "philosopher's stone"

Ω_n을 알면 n-bit thm의 참 거짓을 알 수 있다.

(일반적으로 불가능 by Godel's incompleteness thm).

iii) Ω is algorithmically random.

Let we know Ω_n, where $|p| = n$.

To know the program p halts,

We run all programs p' with $|p'| \leq |p|$ parallelly

Note that we know how many (p')'s are halts.

We can run them until ~~the halts~~ the # of halted p''s are the same with the number.

Thm $\exists c$ s.t

$$K(\Omega_n) > n - c. \ \forall n$$

Pf

Consider the following program :

let $a := \omega_1 \cdots \omega_n$.

find x_0 s.t. $K(x_0) > n$

the smallest

print x_0.

This program has a complexity $K(\Omega_n) + c \overset{\text{def of } x_0}{\geq} K(x_0) > n$

$$\Rightarrow K(\Omega_n) > n - c.$$