Information Theory & Coding

Oct 6th 2020

Yesterday:

- $$D(p\|q) = \sum_{u \in \mathcal{U}}' p(u) \log \frac{p(u)}{q(u)}$$

- $D(\ \ ) \geq 0 \quad = 0$ iff $p = q$

- $I(U;V) = D(p_{UV} \| p_U p_V)$

- $$\Pr(\underbrace{U_1 \ldots U_n}_{\text{iid} \sim q} \in T(p,\varepsilon,n)) \doteq 2^{-n[D(p\|q) + O(\varepsilon)]}$$

$\mathcal{U}$ = alphabet suppose $|\mathcal{U}| = K$

Let $q(u) = \frac{1}{K} \quad u \in \mathcal{U}$

$$D(p\|q) = \sum' p(u) \log \frac{p(u)}{1/K} = -H(p) + (\log K) \underbrace{\sum p(u)}_{=1}$$

$$= (\log K) - H(p).$$

$\Rightarrow$ <u>Lemma</u>: $H(U) \leq \log|\mathcal{U}|$.

$=$ iff U is uniformly distrib on $\mathcal{U}$.

Also yesterday: we saw a "correspondence" between u.d. codes & probability distributions:

I.e.: given a u.d. c, $c : \mathcal{U} \to \{0,1\}^*$

Set: $\begin{cases} q(u) = 2^{-\text{length}(c(u))} \\ q(u_0) = 1 - KS(c) \qquad u_0 \notin \mathcal{U} \end{cases}$

Conversely, given q a distrib on $\mathcal{U}$ set

$$\text{length}\, c(u) = \lceil -\log q(u) \rceil$$

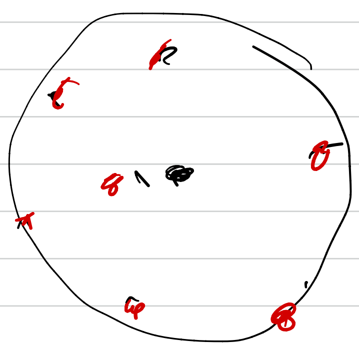
$$\underbrace{E[\text{length}\, c(U)] - H(U)}_{\sim p} \approx D(p\|q).$$

Universal source coding:

Suppose we know that distrib. of U belongs to a set $\mathcal{P}$ of distributions.

A measure of a code $c \longleftrightarrow q$ could be

$$\sup\{D(p\|q) : p \in \mathcal{P}\}$$



$\bullet \in \mathcal{P}$

Finding the q to minimize

Lempel-Ziv universal coding. (late 70's)

Algorithm; explained by Example:

$\boxed{\mathcal{U} = \{a, b, c\}}$.

To describe the sequence

a|ab|aa|b|c|ca|c - - - - - -

Start with a dictionary: $\mathcal{D}_0 = \{a, b, c\}$.

$a \to 00$, $b \to 01$, $c \to 10$

output: 00 00 1000 . . . .

$\mathcal{D}_1 = \{aa, \cancel{ab}, ac, b, c\}$

000 001 010 011, 100

$\mathcal{D}_2 = \{aa, aba, abb, abc, ac, b, c\}$

$aa \to 000$, $c \to 110$

also: given $\mathcal{U} = \{a, b, c\}$ & lexicographic order we can "invert" the operation, to figure out $u_1 u_2 u_3$ — from the binary output of LZ.

We will show LZ does well by showing that it competes well against a strong group of competitors.

Idea: given $u_1 u_2 u_3$ - - - - -

and a Finite State Machine:

Def: a finite state Machine is described by

a set $\mathcal{S}$ of states $|\mathcal{S}| < \infty$.

a $s_0 \in \mathcal{S}$ (starting state)

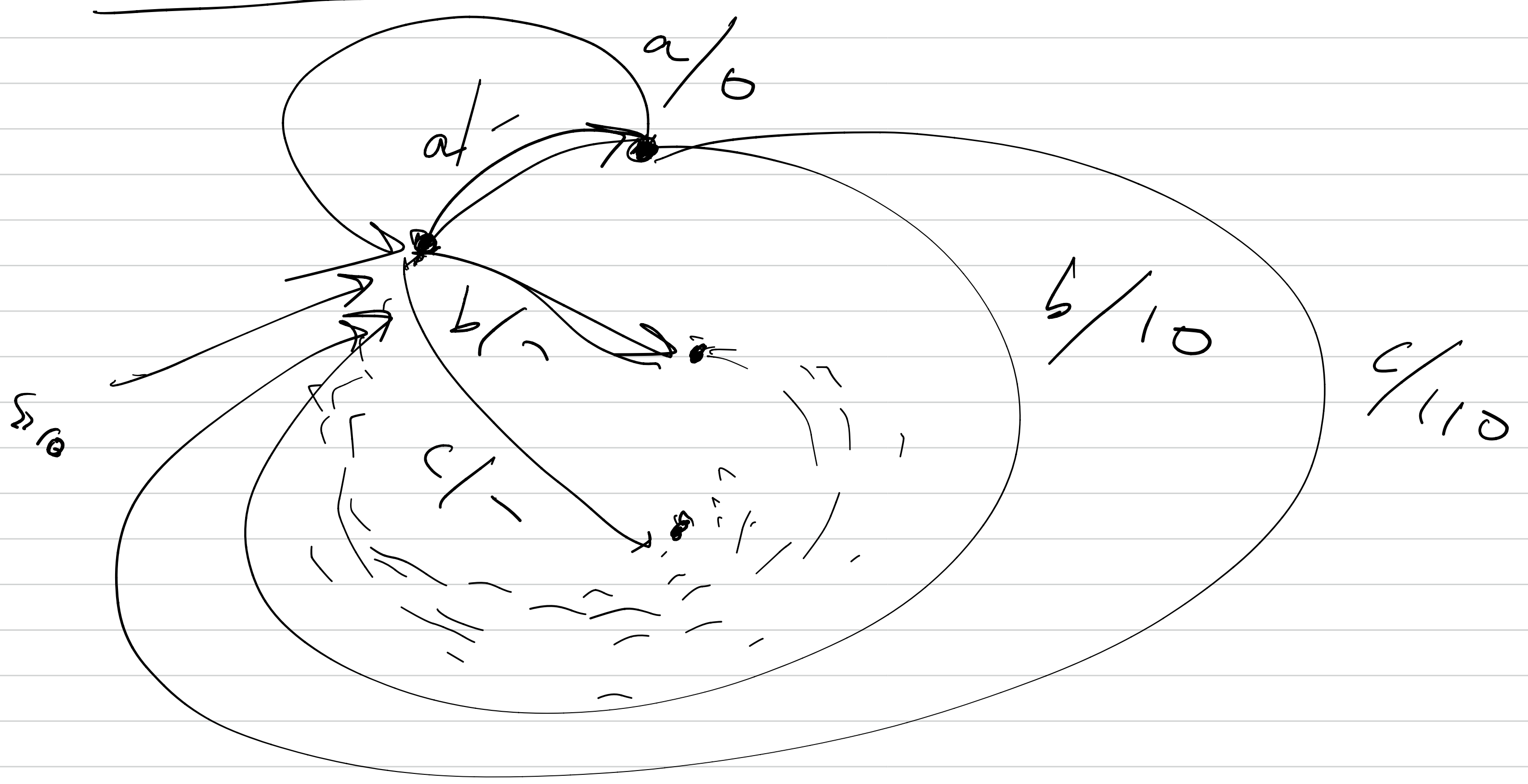
a nextstate function:

$$g : \mathcal{S} \times \mathcal{U} \to \mathcal{S} \qquad s_{i+1} = g(s_i, u_{i+1}) \quad i \geq 0$$

A output function

$$f : \mathcal{S} \times \mathcal{U} \to \{0,1\}^* \qquad y_{i+1} = f(s_i, u_{i+1}) \quad i \geq 0.$$

Example: $\mathcal{U} = \{a, b, c\}$



$\equiv$

aa → 0
ab → 10
ac → 110
ba → .
bb → .
bc → .
ca → .
cb → .
cc → .

$$\frac{1}{n}\left(H(u_1 \ldots u_n) + 1\right)$$

We want "invertible" F.S.M.s. namely

$$\boxed{\begin{array}{l} \text{if } \underline{u_1u_2u_3}\cdots \neq \underline{u_1'u_2'u_3'}\cdots \\ \text{then } \underline{\text{output}(u_1u_2\ldots)} \neq \text{output}(u_1'u_2'\ldots) \end{array}}$$

IL (information lossless) machines:

first extend the domain of f & g, to

define $g(s, u_1u_2) = g(g(s,u_1), u_2)$

$$g(s, u_1 \ldots u_n) = g(g(\ldots g(g(s,u_1),u_2)\ldots u_{n-1}), u_n)$$

$$f(s, u_1u_2) = f(s,u_1)\,f(g(s,u_1), u_2)$$

$f(s, u_1 \ldots u_n)$ = similarly as the totality of the output produced by the machine when it is fed $u_1 \ldots u_n$ & starts at s.

(Also assume that all elements of $\mathcal{S}$ are reachable from s_0.)

IL machine means this:

$$\forall s \in \mathcal{S} \quad \forall \underline{u_1 \ldots u_n} \neq \underline{u_1' \ldots u_m'}$$

either $g(s, u_1 \ldots u_n) \neq g(s, u_1' \ldots u_m')$

or $\underline{f(s, u_1 \ldots u_n) \neq f(s, u_1' \ldots u_m')}$.

Clearly not-IL $\Rightarrow$ not-invertible

$\equiv$ invertible $\Rightarrow$ IL

Def: given a machine M and an infinite sequence $u_1 u_2 \ldots \ldots$ let

$$\rho(M, \vec{u}) = \limsup_{n\to\infty} \frac{1}{n} \operatorname{len}(\text{output}(M, u_1 \ldots u_n))$$

$$\limsup_{n\to\infty} a_n = a \equiv \forall \varepsilon > 0, \exists n_0 \text{ s.t } \forall n > n_0$$
$$a_n < a + \varepsilon \;\&\; \exists n_1 > n_0$$
$$\text{s.t } a_{n_1} > a - \varepsilon$$

$$= \lim_{n\to\infty} \left(\sup_{m \geq n} a_m \right)$$

n

Def: given an positive integer m, define

$$\rho(m, u_1 u_2 u_3 \ldots) = \min\{\rho(M, u_1 u_2 \ldots) : M, \text{IL}, \text{ with } \leq m \text{ states}\}$$

Def: $\rho(u_1 u_2 \ldots) = \lim_{m\to\infty} \rho(m, u_1 u_2 \ldots)$

finite-state compressibility ~~of the sequence~~ $u_1 u_2 \ldots$

We will show :

$$\rho(lz, u_1 u_2 \ldots) \leq \rho(u_1 u_2 \ldots)$$

$\forall u_1 u_2 \ldots$