Self-delimiting Kolmogorov Complexity

Recall C is not subadditive $O(1) + C(x) + C(y) > C(xy)$

(problem)

Not a problem, but a feature:

C is not monotone over prefixes:

say: S_n is an incompressible string — stands for number n.

say has length m.

~~Consider 0^{m+1}~~ — stands for some number n'

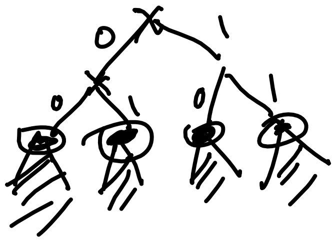
Consider 0^n and $0^{n'}$. $C(0^n) \approx C(n) = m$ but $C(0^{n'}) \leq C(n') = \log(m+1)$. $n' > n$.

$C(x)$ $C(y)$

p_x p_y

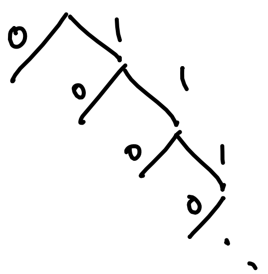
Prefix Codes

A set of strings P is a prefix code (equivalently, a prefix-free code, almost equivalently, instantaneous code) if for every x in P, no proper prefix or proper extension of x can be in P.



e.g. a prefix code can be $\{00, 01, 10, 11\}$.

an infinite prefix-free code:



$1^*0 = \{0, 10, 110, 1110, \ldots\}$

is a canonical infinite prefix-free language

$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots$

≤ 1

if x & $y \in P$ prefix code, then xy can be uniquely parsed into x & y.

Suppose Turing machines are encoded using strings from some prefix code $\underset{\sim}{P}$:

Let $\hat{p}_x$ be the shortest program for x, $\hat{p}_y$ --- y from $\underset{\sim}{P}$.

$\hat{p}_x \hat{p}_y$

Lemma (Kraft Inequality)

Let P be any prefix code. Then

$$\sum_{x \in P} \frac{1}{2^{|x|}} \leq 1.$$