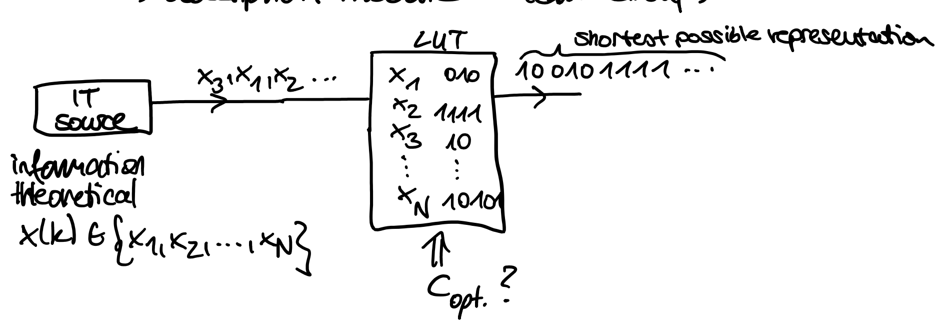
10. lecture

Freitag, 20. Oktober 2023 09:56

6pm + laptop mitbringen
→ description moodle (TeamsGroup)



Memoryless & stacionary source model

$$p(x(k) = x_i \mid \cancel{x(k-1) = x_j \dots} \; x(0) = x_M) = p(x(k) = x_i) = p_i(k) = p_i$$

$p_1 \; p_2 \cdots p_N \to p(x)$
$x_1 \; x_2 \cdots x_N \to x$
$\bar{c}_1 \; \bar{c}_2 \cdots \bar{c}_N \to \bar{c}(x)$
$\ell_1 \; \ell_2 \cdots \ell_N \to \ell(x)$ length of codeword

Average codelength: $L := E(\ell(x)) = \sum_x p(x)\,\ell(x)$

$C_{opt}: \min_c L$

$H(x) = \sum_x p(x) \operatorname{ld} \frac{1}{p(x)}$ $\qquad 0 \leq H(x) \leq \operatorname{ld} N$

average inform. provided by source = Entropie

Kraft inequality: $\sum_x 2^{-\ell(x)} \leq 1$ (binary tree – prefix free)

$H(x) \leq L$

Shannon – Fano codes:

$$H(x) = \sum_x p(x) \operatorname{ld} \frac{1}{p(x)} \leq \sum_x p(x) \left\lceil \operatorname{ld} \frac{1}{p(x)} \right\rceil$$

Algo: given $p(x)$ – source distribution $\mapsto \ell(x) := \left\lceil \operatorname{ld} \frac{1}{p(x)} \right\rceil$
→ binary tree → $\bar{c}(x)$ → LUT

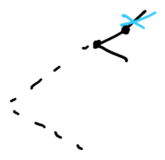
$\sum_x 2^{-\ell(x)} \leq 1$ (prefix free ✓)

$H(x) \leq L_{SF} \leq H(x) + 1$

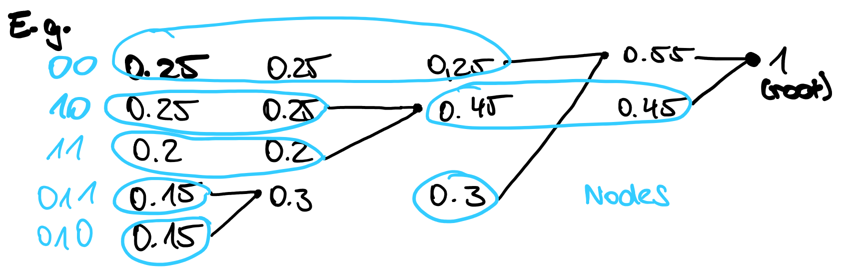
Huffman coding

$p_1 > p_2 > \dots > p_{N-1} > p_N$

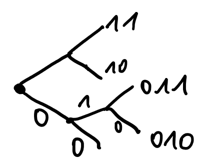
Criteria for optimality → if $p_i > p_j$ then $\ell_i < \ell_j$
→ $\ell_{N-1} = \ell_N$



E.g.

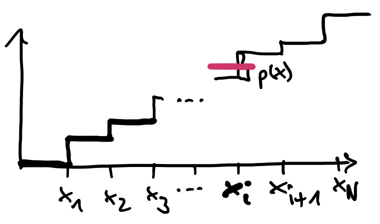


1. two smallest 2. two smallest ↯ throwback → time window for probabilities
↳ in every time window calculate the code new



Shannon – Fano – Elias coding

$F(x_i) = \sum_{j=1}^{i} p(x_j)$ $\qquad \tilde{F}(x_i) = \sum_{j=1}^{i-1} p(x_j) + \frac{1}{2} p(x_i)$



$\tilde{F}(x_i) - \lfloor \tilde{F}(x_i) \rfloor_{\ell(x)} \leq \frac{1}{2^{\ell(x)}}$

$\ell(x) := \left\lceil \operatorname{ld} \frac{1}{p(x)} \right\rceil + 1 \quad \to \quad \frac{1}{2^{\ell(x)}} = \frac{1}{2^{\lceil \operatorname{ld} \frac{1}{p(x)} \rceil + 1}} \leq \frac{p(x)}{2}$

$$\sum_x 2^{-\ell(x)} = \sum_x 2^{-\lceil \operatorname{ld} \frac{1}{p(x)} \rceil - 1} \leq \sum_x 2^{-\operatorname{ld} \frac{1}{p(x)} - 1} = \sum_x \frac{1}{2} \sum_x 2^{-\operatorname{ld} \frac{1}{p(x)}}$$
$$= \sum_x \frac{1}{2} p(x)$$
$$= \frac{1}{2} \leq 1$$

$$L = \sum_x \ell(x) p(x) = \sum_x \left(\left\lceil \operatorname{ld} \frac{1}{p(x)} \right\rceil + 1 \right) p(x) \leq \sum_x \left\{ \left(\operatorname{ld} \frac{1}{p(x)} + 1 \right) + 1 \right\} p(x)$$
$$= H(x) + 2$$

$$H(x) \leq L_{SFE} \leq H(x) + 2$$

Algo: given $p(x_i)$, $i = 1, \dots, N \to \tilde{F}(x_i) = \sum_{j=1}^{i-1} p(x_j) + \frac{1}{2} p(x_i)$
$\to P(x_i) = \left\lceil \operatorname{ld} + \frac{1}{p(x_i)} \right\rceil + 1 \to \lfloor \tilde{F}(x_i) \rfloor_{\ell(x)} \to \bar{c}(x_i)$

(no binary tree, no searching for two minimal propability)
↳ price: average L größer dafür faster

→ small improvement in „2" matters a lot in dataspeed

Huffman 372,8 Mbps binary tree
SF
SFE ↓ higher faster als Huffman

⇒ parametric methods: know the propability distributions before