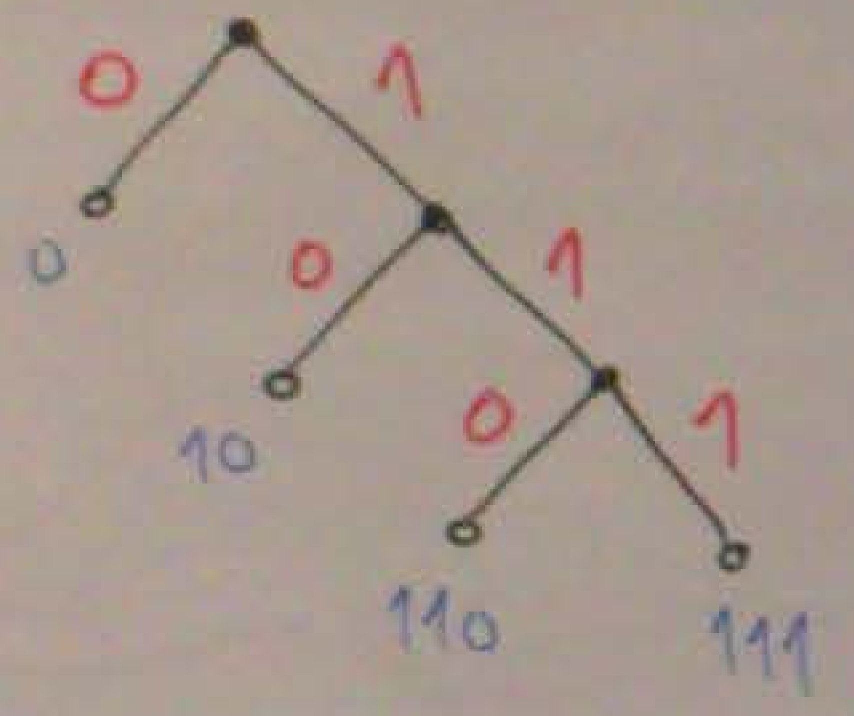
Huffmann coding

- prefix-free code

  symbol length is given a-priori

  (see Kraft ineq.)

Ex: $l_i \in \{1, 2, 3, 3\}$



Q: why $\{1, 2, 3, 3\}$ and not $\{2, 2, 2, 2\}$?

Q: which source symbol shall be assigned when?

↳ this should reflect the entropy (or, p_i)

Ex: $A \to p(A) = 0.5$

$B \to p(A) = 0.25$

$C \to p(C) = 0.125$

$D \to p(D) = 0.125$

descending order

0.5 | 0.5 | $\genfrac{}{}{0pt}{}{0}{1}$

0.25 $\genfrac{}{}{0pt}{}{0}{1}$ | 0.5

0.25 $\genfrac{}{}{0pt}{}{0}{1}$

$p(A) = \frac{1}{2^1}$

$p(B) = \frac{1}{2^2}$

$p(D) = p(C) = \frac{1}{2^3}$

↳ optimal!

Huff. c. works well only for these!!

Ex: $p(A) \to 0.7$

$p(B) \to 0.2$

$p(C) \to 0.09$

$p(D) \to 0.01$

$p(C), p(D) \to 0.1$; $0.1, p(B) \to 0.3$

Huffmann coding

• problems:

a) optimality

b) 2-pass approach

pass #1: determine p_i (source symbol probabilities)

pass #2: encode

⇒ adaptive Huffman coding

start with a-priori distribution and upate it on-line

→ re-construction of Huffman tree

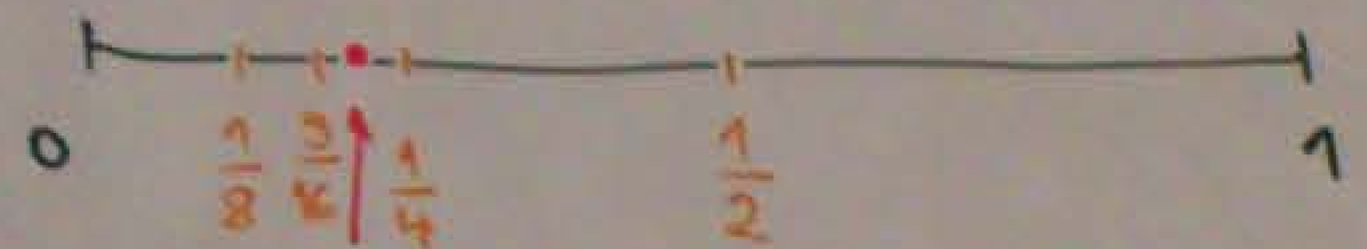
c) it is symbol-based

in English: 'the', 'a', 'is', 'qu' are very frequent

Arithmetic coding

arbitrary precision arithmetic

Idea: Represent the message as a rational number and represent this rational number in binary arithmetic on [0,1]



$$\frac{11}{16} = \frac{1}{2} + \frac{0}{4} + \frac{1}{8} + \frac{1}{16}$$

0.1011

$0.1_2 = 2^{-1} = \frac{1}{2}$, $0.01_2 = 2^{-2} = \frac{1}{4}$

a) $\frac{a}{b}$ ... a fraction representing the message

b) look for an interval in [0,1] where $\frac{a}{b}$ lies by binary search (bisection)

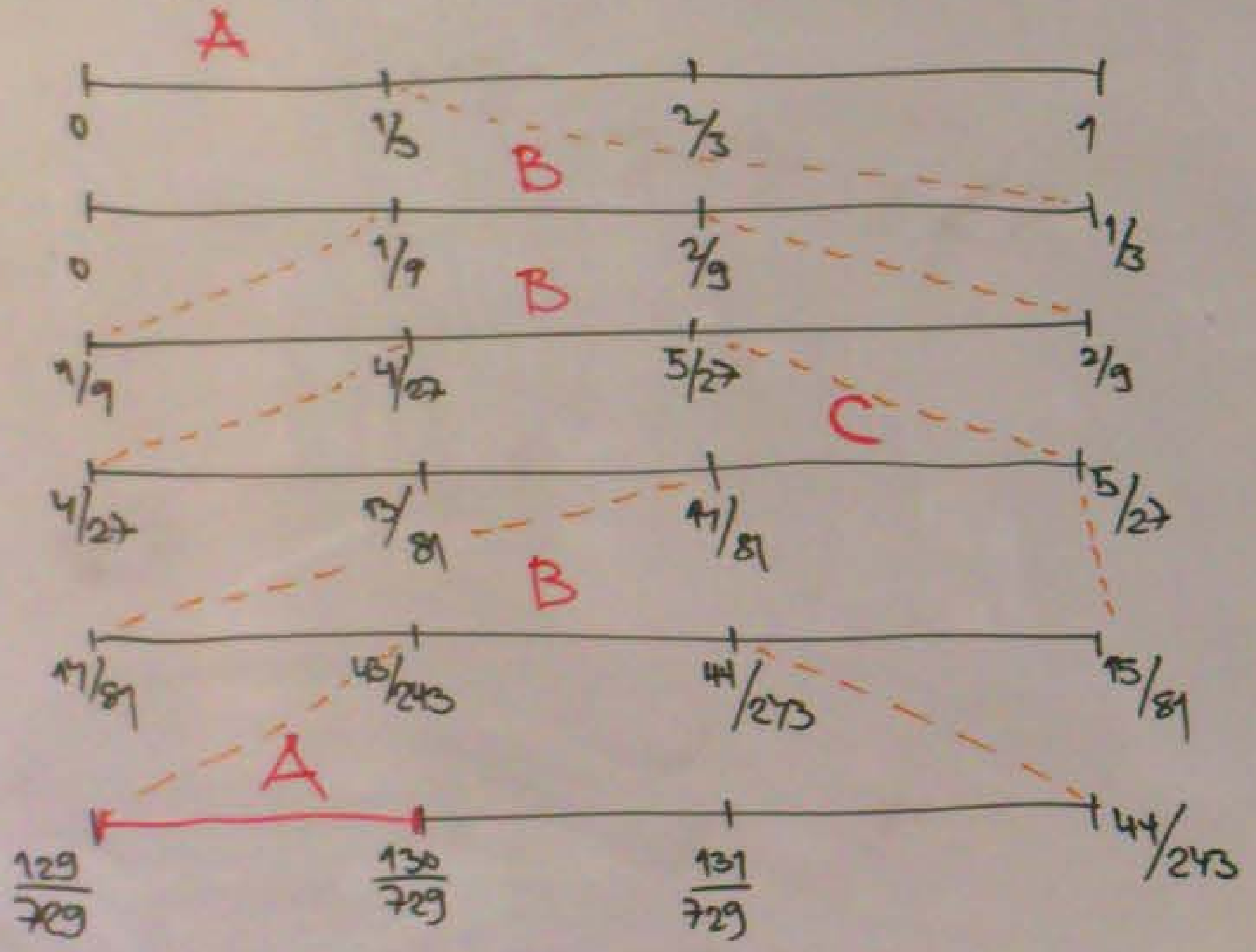
→ „bracket"

Ex: $(0{,}1)_{10} \equiv (0.0001)_2 \in [\frac{1}{16} | \frac{2}{16}]$

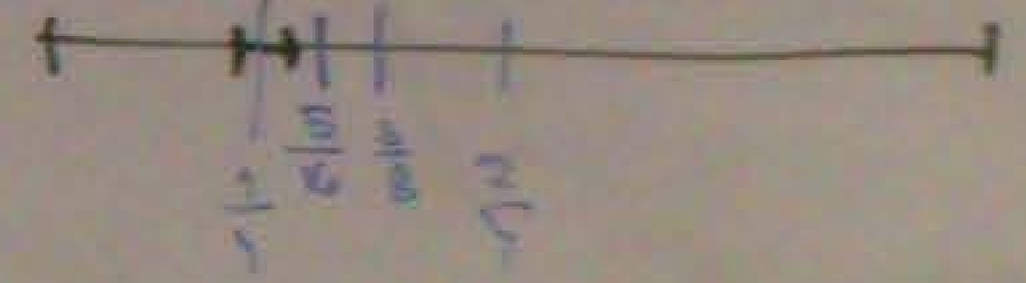
and $(0.2)_{10} \in [\frac{2}{16} | \frac{3}{16}]$

Ex: $p(A) = p(B) = p(C) = 1/3$

Arithmetic code for ABBCBA



this is the representation of ABBCBA



now: find a binary fraction $\frac{x}{2^k}$ such that

$$\frac{x}{2^k} \in \left(\frac{129}{729} , \frac{130}{729} \right)$$

! the 'shortest' one !

the binary string representing the fraction is the encoded message!

$$\frac{129}{729} = 0.17695 < \frac{x}{2^k} < 0.17832 = \frac{130}{729}$$

$k=1$: $\frac{1}{2} > 0.17$ ✗

$k=2$: $\frac{1}{4} > 0.17$ ✗

⋮

$k=7$: $\frac{22}{128} < 0.176$, $\frac{23}{128} > 0.178$ ✗

$k=8$: $\frac{45}{256} < 0.176$, $\frac{46}{256} > 0.178$ ✗

$k=9$: $\frac{91}{512} > 0.176$, $\frac{91}{512} < 0.178$ ✓

binary rep. of $\frac{91}{512} = \frac{1}{8} + \frac{1}{32} + \frac{1}{64} + \frac{1}{256} + \frac{1}{512} \Rightarrow 0.\boxed{001011011}$