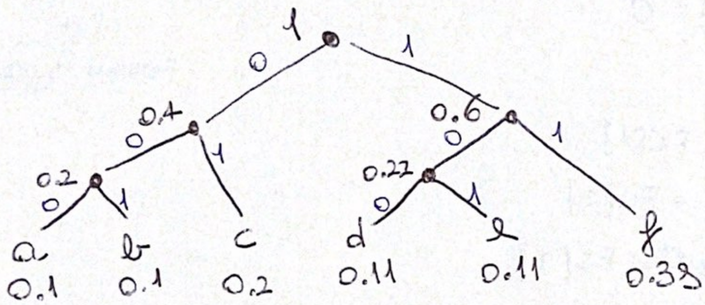
CANONICAL HUFFMAN (1)

SYMB	a	b	c	d	e	f
PROB	0.1	0.1	0.2	0.11	0.11	0.38

a) COMPUTE THE HUFFMAN CODE



a = 000
b = 001
c = 01
d = 100
e = 101
f = 11

b) COMPUTE THE CANONICAL HUFFMAN CODE

STEP 0. COMPUTE CODEWORD LENGHT ∀ SYMBOL

STEP 1. CONSTRUCT NUM ARRAY. NUM[l] = # SYMBOLS W. CODEWORD LENGHT l

l	1	2	3
NUM	0	2	4

STEP 2. CONSTRUCT SYMBOL TABLE

l	SYMBOL
1	
2	c, f
3	a, b, d, e

0 1 2 3

STEP 3. COMPUTE THE "FIRST-CODEWORD" ARRAY FC

$FC[l_{MAX}] = 0$

$FC[l] = (FC[l+1] + NUM[l+1]) / 2$

l	1	2	3
FC	2	2	0

c) ENCODE T = f a d e

OFFSET INTO SYMBOL LIST

f	a	d	e
FC[2]+1 IN 2 BITS	FC[3]+0 IN 3 BITS	FC[3]+2 IN 3 BITS	FC[3]+3 IN 3 BITS
11	000	010	011

d) DECODE THE FIRST 3 SYMBOLS OF

1 0 0 0 1 0 1 1 0 1 1 1 0 0 ...

1ST SYMBOL

- $\ell = 1 \quad v = (1)_2 < 2 = FC[1]$
- $\ell = 2 \quad v = (10)_2 \geq 2 = FC[2]$

OUT $SYMB[2, 2-2] = c$

2ND SYMBOL

- $\ell = 1 \quad v = (0)_2 < 2 = FC[1]$
- $\ell = 2 \quad v = (00)_2 < 2 = FC[2]$
- $\ell = 3 \quad v = (001)_2 \geq 0 = FC[3]$

OUT $SYMB[3, 1-0] = b$

3RD SYMBOL

- $\ell = 1 \quad v = (0)_2 < 2 = FC[1]$
- $\ell = 2 \quad v = (01)_2 < 2 = FC[2]$
- $\ell = 3 \quad v = (011)_2 \geq 0 = FC[3]$

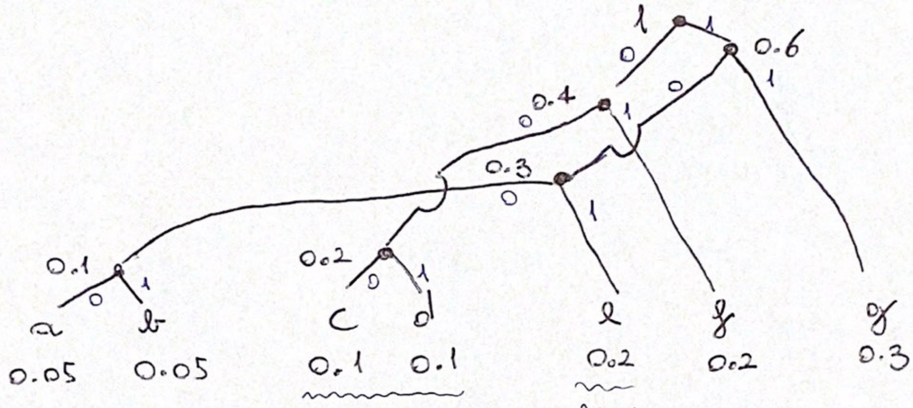
OUT $SYMB[3, 3-0] = \ell$

```
v = NEXTBIT()
l = 1
WHILE v < FC[l]
    v = 2v + NEXTBIT()
    l++
RETURN SYMB[l, v-FC[l]]
```

CANONICAL HUFFMAN ②

SYMB	a	b	c	d	e	f	g
PROB	0.05	0.05	0.1	0.1	0.2	0.2	0.3

a) COMPUTE THE HUFFMAN CODE



a = 1000
b = 1001
c = 000
d = 001
e = 101
f = 01
g = 11

OLDEST STRATEGY
PICK NODES THAT HAVE BEEN MERGED FARTHEST IN THE PAST

AGAIN

b) COMPUTE THE CANONICAL HUFFMAN

l	1	2	3	4
NUM	0	2	3	2

l	SYMBOL
1	
2	f, g
3	c, d, e
4	a, b

0 1 2

l	1	2	3	4
FC	2	2	1	0

c) ENCODE $T = g\,c\,d$

g: FC[2]+1 IN 2 BITS → 11

c: FC[3]+0 IN 3 BITS → 001

d: FC[3]+1 → 010

d) DECODE FIRST 3 SYMBOLS OF 0 0 0 1 0 1 1 0 1 0 0 1 0 ...

1ST SYMBOL
- $l=1$ $v=(0)_2 < 2 = FC[1]$
- $l=2$ $v=(00)_2 < 2 = FC[2]$
- $l=3$ $v=(000)_2 < 1 = FC[3]$
- $l=4$ $v=(0001)_2 \geq 0 = FC[4]$

OUT SYMB[4, 1-0] = b

2ND SYMBOL
- $l=1$ $v=(0)_2 < 2 = FC[1]$
- $l=2$ $v=(01)_2 < 2 = FC[2]$
- $l=3$ $v=(011)_2 \geq 1 = FC[3]$

OUT SYMB[3, 3-1] = e

3RD SYMB
- $l=1$ $v=(0)_2 < 2 = FC[$
- $l=2$ $v=(01)_2 < 2$
- $l=3$ $v=(010)_2 \geq 1 = F$

OUT SYMB[3, 2-1] = d

ARITHMETIC CODE

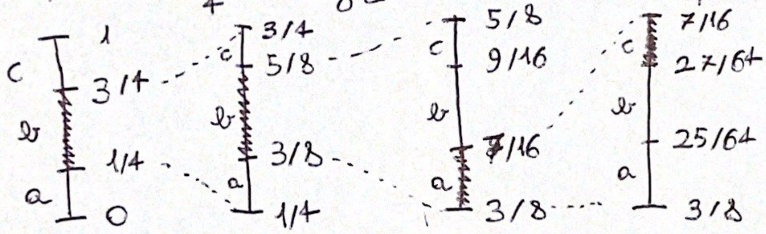
① ENCODE THE TEXT $T = b\,b\,a\,c$ USING AS PROBABILITIES OF THE SYMBOLS THEIR FREQUENCY IN T

STEP 0: COMPUTE PROB AND CUMULATIVE PROBABILITIES

$P[a] = \frac{1}{4}$ $\qquad f_a = 0$

$P[b] = \cancel{\frac{2}{4}}\ \frac{1}{2}$ $\qquad f_b = P[a] = \frac{1}{4}$

$P[c] = \frac{1}{4}$ $\qquad f_c = P[a] + P[b] = \frac{3}{4}$



STEP 1: DO |T| STEPS OF THE AC-CODING ALGORITHM

- $s_0 = 1$ ← STARTING DOWN
  $l_0 = 0$ ← LENGHT
- $s_1 = s_0 \cdot P[b] = 1 \cdot \frac{1}{2} = \frac{1}{2}$ $\rightsquigarrow \left[\frac{1}{4}, \frac{1}{4} + \frac{1}{2}\right) = \left[\frac{1}{4}, \frac{3}{4}\right)$
  $l_1 = l_0 + s_0 f_b = 0 + 1 \cdot \frac{1}{4} = \frac{1}{4}$
- $s_2 = s_1 \cdot P[b] = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$ $\rightsquigarrow \left[\frac{3}{8}, \frac{3}{8} + \frac{1}{4}\right) = \frac{5}{8}$
  $l_2 = l_1 + s_1 f_b = \frac{1}{4} + \frac{1}{2}\cdot\frac{1}{4} = \frac{1}{4} + \frac{2}{16} = \frac{3}{8}$
- $s_3 = s_2 \cdot P[a] = \frac{1}{4} \cdot \frac{1}{4} = \frac{1}{16}$ $\rightsquigarrow \left[\frac{3}{8}, \frac{3}{8} + \frac{1}{16}\right) = \left[\frac{3}{8}, \frac{7}{16}\right)$
  $l_3 = l_2 + s_2 f_a = \frac{3}{8} + \frac{1}{4} \cdot 0 = \frac{3}{8}$
- $s_4 = s_3 \cdot P[c] = \frac{1}{16} \cdot \frac{1}{4} = \frac{1}{64}$ $\rightsquigarrow \left[\frac{27}{64}, \frac{27}{64} + \frac{1}{64}\right) = \left[\frac{27}{64}, \cancel{\frac{28}{64}}\ \frac{7}{16}\right)$
  $l_4 = l_3 + s_3 f_c = \frac{3}{8} + \frac{1}{16} \cdot \frac{3}{4} = \frac{27}{64}$

STEP 2: CHOOSE A NUMBER INSIDE THE LAST INTERVAL W. THE SMALLEST BIT-LENGHT

$\#BITS = \left\lceil \log_2 \frac{2}{s} \right\rceil = \left\lceil \log_2 \frac{2}{1/64} \right\rceil = \log_2 128 = 7$

NUMBER TO CODE $= l + \frac{s}{2} = \frac{27}{64} + \frac{1}{128} = \frac{55}{128}$ $\rightsquigarrow$ WRITE DOWN IMMEDIATELY IN 7 BITS

- $2 \cdot \frac{55}{128} = \frac{55}{64} < 1 \rightarrow$ OUTPUT 0
- $2 \cdot \frac{55}{64} = \frac{55}{32} \geq 1 \rightarrow$ OUTPUT 1
- $2 \cdot \left(\frac{55}{32} - 1\right) = 2 \cdot \frac{23}{32} = \frac{23}{16} \geq 1 \rightarrow$ OUTPUT 1
- $2 \cdot \left(\frac{23}{16} - 1\right) = 2 \cdot \frac{7}{16} = \frac{7}{8} < 1 \rightarrow$ OUTPUT 0
- $2 \cdot \frac{7}{8} = \frac{7}{4} \geq 1 \rightarrow$ OUTPUT 1
- $2 \cdot \left(\frac{7}{4} - 1\right) = 2 \cdot \frac{3}{4} = \frac{3}{2} \geq 1 \rightarrow$ OUTPUT 1
- $2 \cdot \left(\frac{3}{2} - 1\right) = 2\left(\frac{1}{2}\right) \geq 1 \rightarrow$ OUTPUT 1

$T \xrightarrow{\text{ARITHM.}} (0110111)_2, 4$ ← |T|

ARITHMETIC CODE ②

GIVEN SYMBOLS AND PROBABILITIES $P[a] = 0.2$, $P[b] = P[c] = 0.4$

COMPRESS THE FIRST 3 SYMBOLS OF $T = abc$ VIA ARITHMETIC CODE

$P[a] = \frac{2}{10} = \frac{1}{5}$ $\qquad P[b] = P[c] = \frac{4}{10} = \frac{2}{5}$

$f_a = 0$ $\qquad f_b = \frac{1}{5}$ $\qquad f_c = \frac{1}{5} + \frac{2}{5} = \frac{3}{5}$

$s_0 = 1$

$l_0 = 0$

- $s_1 = s_0 \cdot P[a] = 1 \cdot \frac{1}{5} = \frac{1}{5}$

  $l_1 = l_0 + s_0 \cdot f_a = 0 + 1 \cdot 0 = 0$
- $s_2 = s_1 \cdot P[b] = \frac{1}{5} \cdot \frac{2}{5} = \frac{2}{25}$

  $l_2 = l_1 + s_1 \cdot f_b = 0 + \frac{1}{5} \cdot \frac{1}{5} = \frac{1}{25}$
- $s_3 = s_2 \cdot P[c] = \frac{2}{25} \cdot \frac{2}{5} = \frac{4}{125}$

  $l_3 = l_2 + s_2 \cdot f_c = \frac{1}{25} + \frac{2}{25} \cdot \frac{3}{5} = \frac{1}{25} + \frac{6}{125} = \frac{5+6}{125} = \frac{11}{125}$

$\# \text{BITS} = \lceil \log \frac{2}{s} \rceil = \lceil \log \frac{2}{4/125} \rceil = \lceil \log \frac{250}{4} \rceil = \lceil \log 62.5 \rceil = 6$

NUMBER TO CODE $l + \frac{s}{2} = \frac{11}{125} + \frac{\cancel{4}^{2}}{125} \cdot \frac{1}{\cancel{2}_1} = \frac{13}{125}$

- $2 \cdot \frac{13}{125} = \frac{26}{125} < 1 \rightarrow$ OUTPUT 0
- $2 \cdot \frac{26}{125} = \frac{52}{125} < 1 \rightarrow$ OUTPUT 0
- $2 \cdot \frac{52}{125} = \frac{104}{125} < 1 \rightarrow$ OUTPUT 0
- $2 \cdot \frac{104}{125} = \frac{208}{125} \geq 1 \rightarrow$ OUTPUT 1
- $2 \cdot \left(\frac{208}{125} - 1\right) = 2 \cdot \left(\frac{83}{125}\right) = \frac{166}{125} \geq 1 \rightarrow$ OUTPUT 1
- $2 \cdot \left(\frac{166}{125} - 1\right) = 2 \cdot \frac{41}{125} = \frac{82}{125} < 1 \rightarrow$ OUTPUT 0

$T \overset{AC}{\rightsquigarrow} (000110)_2, 3 \leftarrow |T|$

ARITHMETIC CODE (3)

GIVEN SYMBOLS AND PROBABILITIES $P[a]=0.5$, $P[b]=P[c]=0.25$

DECOMPRESS THE FIRST 4 SYMBOLS OF THE ARITHMETIC-CODED BIT SEQ 10111

$P[a]=\frac{1}{2}$ $\quad P[b]=\frac{1}{4}$ $\quad P[c]=\frac{1}{4}$

$f_a=0$ $\quad f_b=\frac{1}{2}$ $\quad f_c=\frac{1}{2}+\frac{1}{4}=\frac{3}{4}$

$$v=(0.10111)_2=\frac{1+2+4+16}{2^5}=\frac{23}{32}$$

- 1
- c
- $3/4 = 24/32$
- b
- $1/2 = 16/32$
- a
- 0

$r_0=1$

- $3/4$
- c
- $l_1+r_1\cdot f_c=\frac{1}{2}+\frac{1}{4}\cdot\frac{3}{4}=\frac{11}{16}=\frac{22}{32}$
- b
- $l_1+r_1 f_b=\frac{1}{2}+\frac{1}{4}\cdot\frac{1}{2}=\frac{5}{8}=\frac{20}{32}$
- a
- $1/2=l_1$

$r_1=r_0\cdot P[b]=\frac{1}{4}$

- $3/4$
- c
- $l_2+r_2 f_c=\frac{11}{16}+\frac{1}{16}\cdot\frac{3}{4}=\frac{47}{64}$
- b
- $l_2+r_2\cdot f_b=\frac{11}{16}+\frac{1}{16}\frac{1}{2}=\frac{23}{32}$
- a
- $\frac{11}{16}=l_2$

$r_2=r_1 P[c]=\frac{1}{4}\cdot\frac{1}{4}=\frac{1}{16}$

- $47/64$
- $l_3+r_3 f_c=\frac{23}{32}+\frac{1}{64}\cdot\frac{3}{4}=\frac{187}{256}$
- $l_3+r_3 f_b=\frac{23}{32}+\frac{1}{64}\cdot\frac{1}{2}=\frac{93}{128}$
- $23/32=l_3$

$r_3=r_2 P[b]=\frac{1}{64}$

$$(10111)_2, 4 \xrightarrow{\text{ARITHM DEC}} bcba$$

THMETIC CODE (4)

GIVEN SYMBOLS AND PROBABILITIES $P[a]=0.25$ $P[b]=0.5$ $P[c]=0.25$

DECOMPRESS THE FIRST 3 SYMBOLS OF THE ARITHMETIC-CODED BASE Q 01101

$P[a] = \frac{1}{4}$ $\quad P[b] = \frac{1}{2}$ $\quad P[c] = \frac{1}{4}$

$f_a = 0$ $\quad f_b = \frac{1}{4}$ $\quad f_c = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}$

$x = (0.01101)_2 = \frac{1+4+8}{2^5} = \frac{13}{32}$

1

c

$3/4 = \frac{24}{32}$

b

$1/4 = \frac{8}{32}$

a

$0 = l_0$

$r_0 = 1$

3/4

c

$l_1 + r_1 f_c = \frac{1}{4} + \frac{1}{2} \cdot \frac{3}{4} = \frac{1}{4} + \frac{3}{8} = \frac{5}{8} = \frac{20}{32}$

b

$l_1 + r_1 f_b = \frac{1}{4} + \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{4} + \frac{1}{8} = \frac{3}{8} = \frac{12}{32}$

a

$1/4 = l_1$

$r_1 = r_0 \cdot P[b] = \frac{1}{2}$

5/8

c

$l_2 + r_2 f_c = \frac{3}{8} + \frac{1}{4} \cdot \frac{3}{4} = \frac{3}{8} + \frac{3}{16} = \frac{9}{16} = \frac{18}{32}$

b

$l_2 + r_2 f_b = \frac{3}{8} + \frac{1}{4} \cdot \frac{1}{4} = \frac{3}{8} + \frac{1}{16} = \frac{7}{16} = \frac{14}{32}$

a

$3/8 = l_2$

$r_2 = r_1 \cdot P[b] = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

$(01101)_2, 3 \xrightarrow[\text{DEC}]{\text{ARITHM}} b\,b\,a$