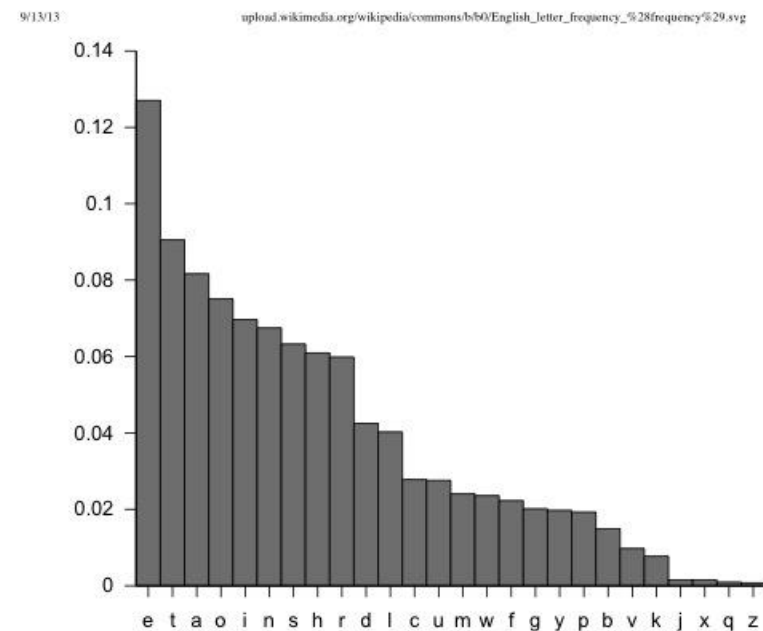
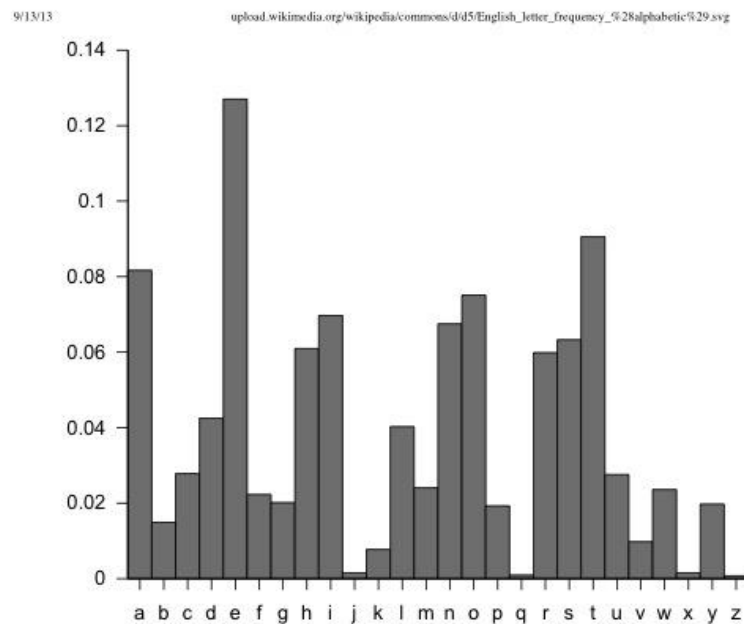


Letter frequency in English

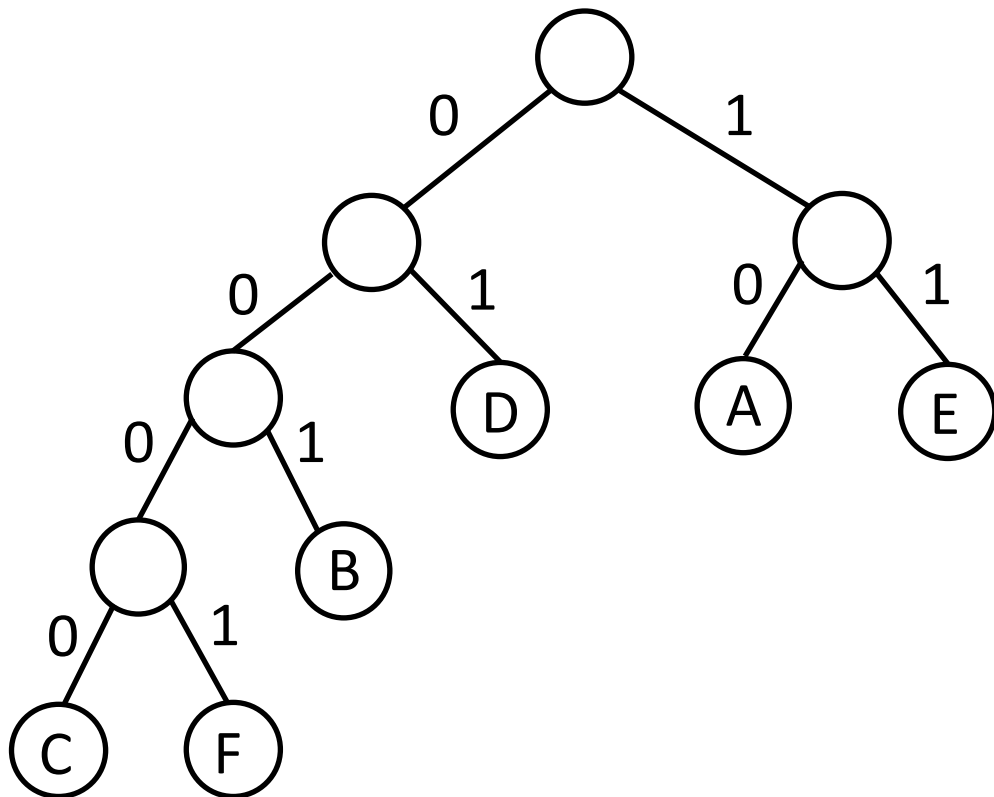
source: Wikipedia

http://en.wikipedia.org/wiki/Letter_frequency



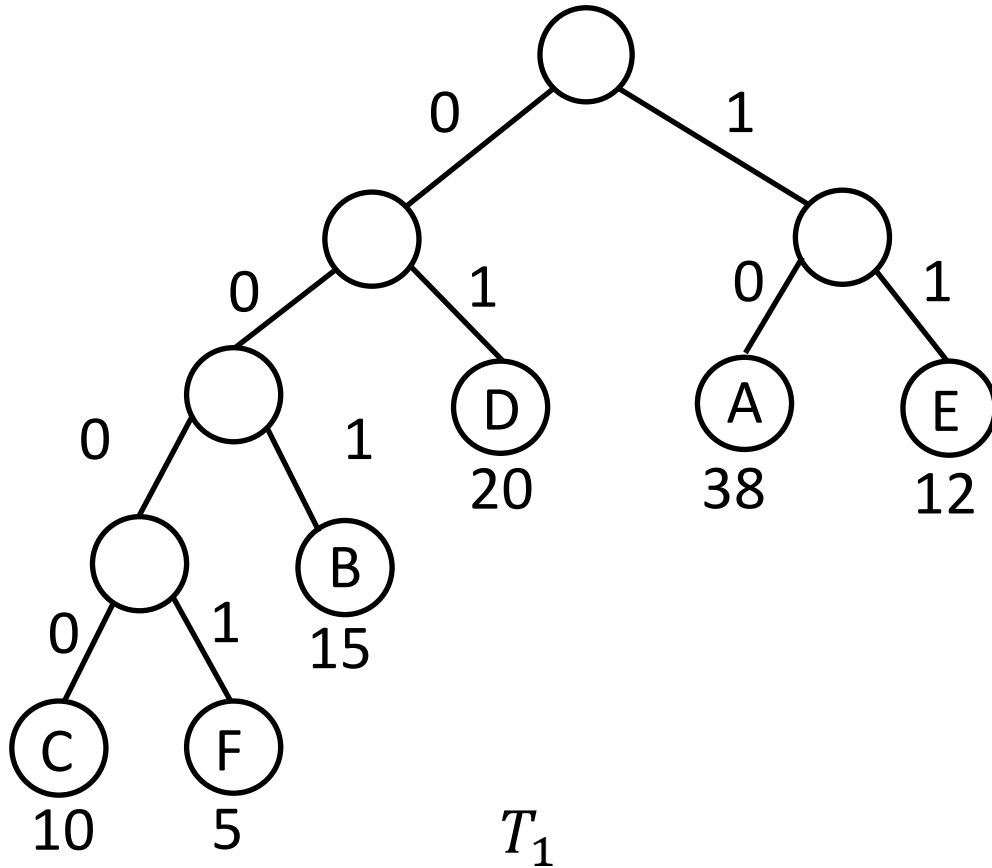
Example of binary tree representation of prefix code

$\Gamma = \{A, B, C, D, E, F\}$



Symbol	Codeword
A	10
B	001
C	0000
D	01
E	11
F	0001

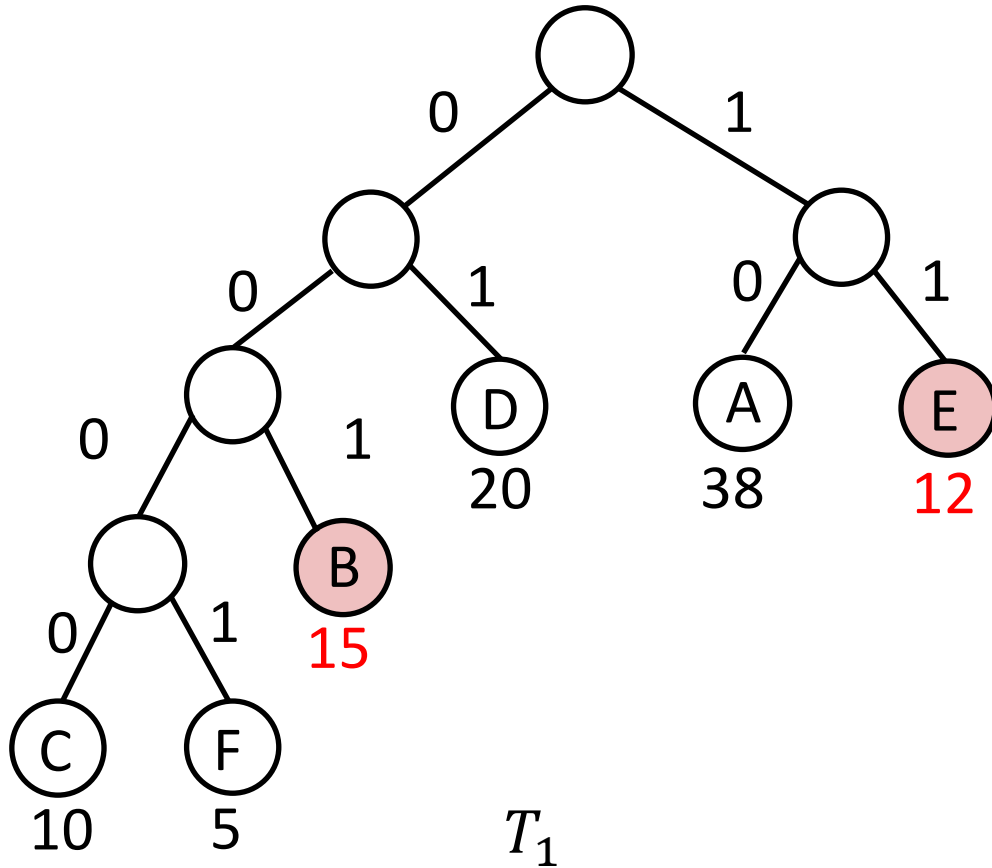
Average depth, given code and frequencies



Symbol	Frequency	Codeword
A	0.38	10
B	0.15	001
C	0.10	0000
D	0.20	01
E	0.12	11
F	0.05	0001

$$AD(T_1) = 0.10 \times 4 + 0.05 \times 4 + 0.15 \times 3 + 0.20 \times 2 + 0.38 \times 2 + 0.12 \times 2 = 2.45$$

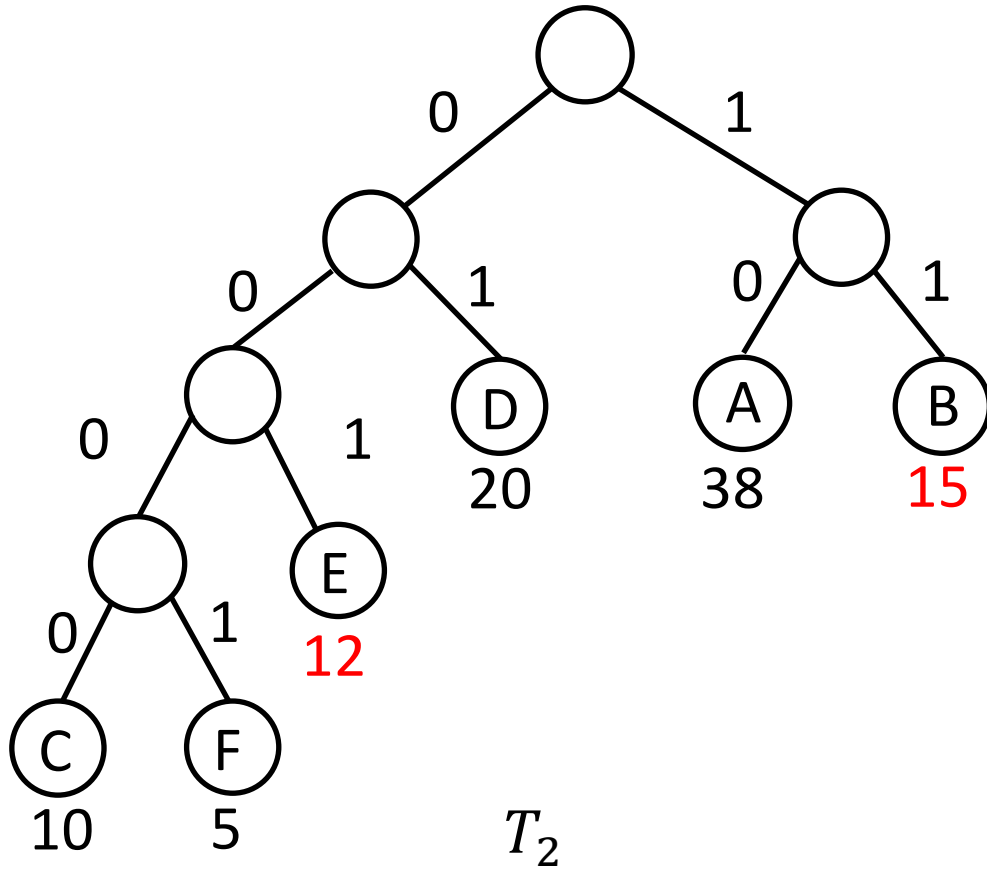
Average depth, given code and frequencies



Symbol	Frequency	Codeword
A	0.38	10
B	0.15	001
C	0.10	0000
D	0.20	01
E	0.12	11
F	0.05	0001

$$AD(T_1) = 0.10 \times 4 + 0.05 \times 4 + 0.15 \times 3 + 0.20 \times 2 + 0.38 \times 2 + 0.12 \times 2 = 2.45$$

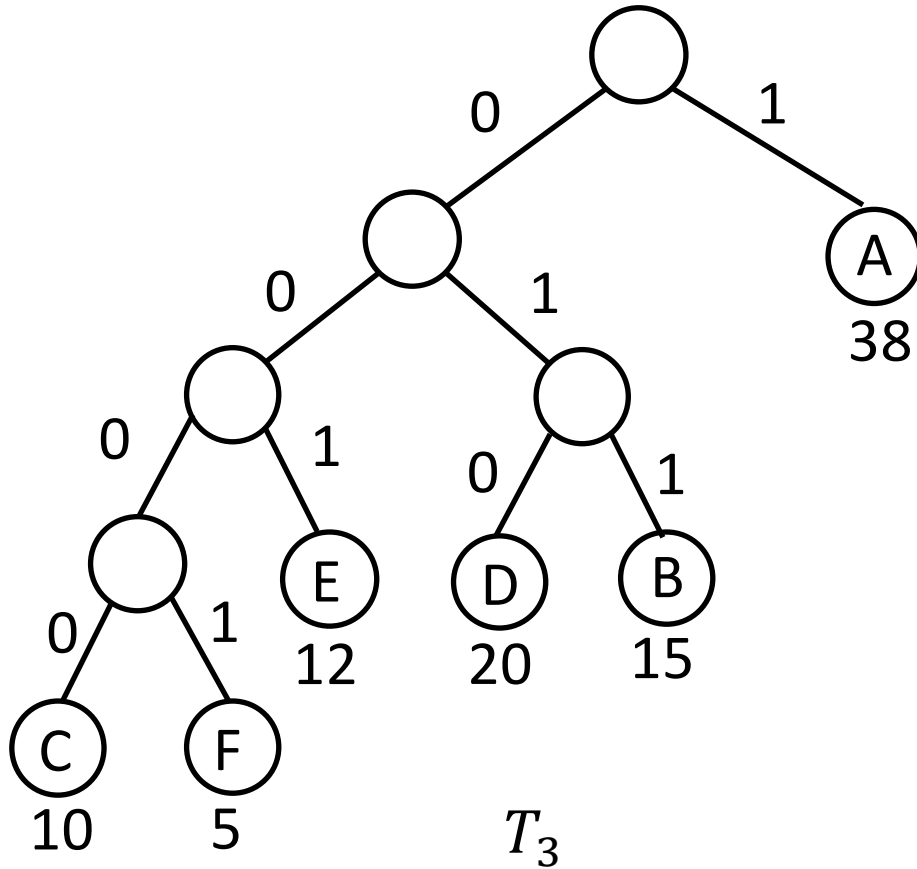
Improved code



Symbol	Frequency	Codeword
A	0.38	10
B	0.15	11
C	0.10	0000
D	0.20	01
E	0.12	001
F	0.05	0001

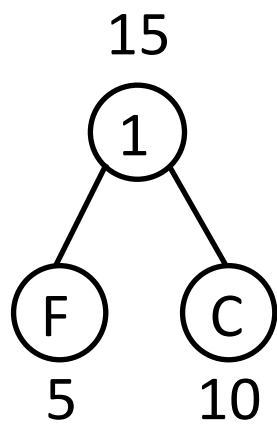
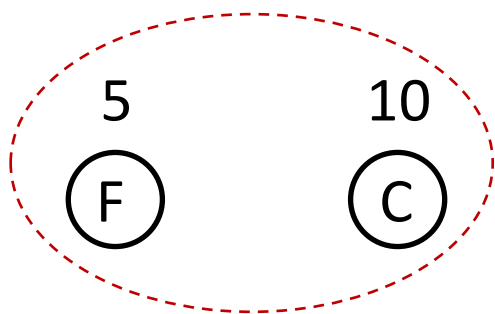
$$AD(T_2) = 0.10 \times 4 + 0.05 \times 4 + 0.12 \times 3 + 0.20 \times 2 + 0.38 \times 2 + 0.15 \times 2 = 2.42$$

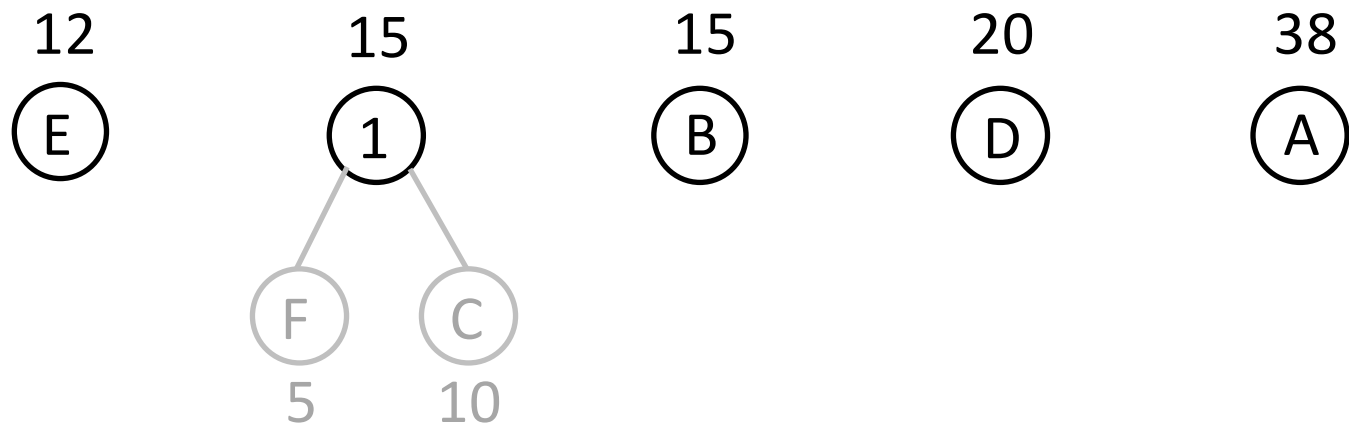
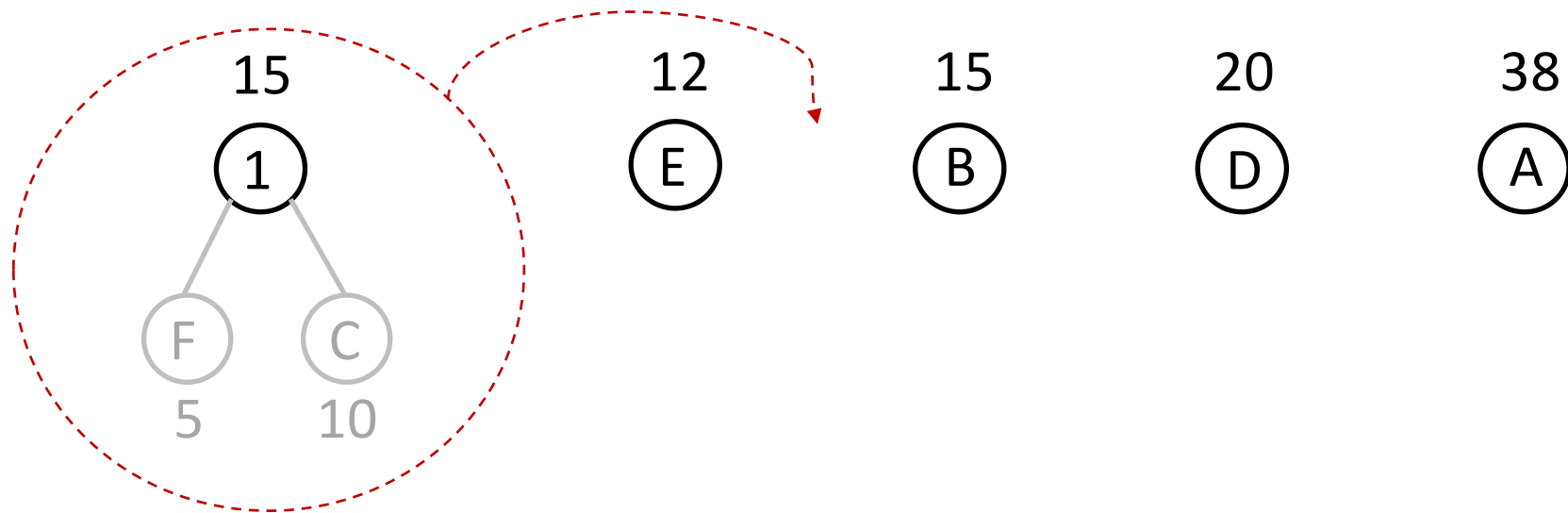
Optimal code

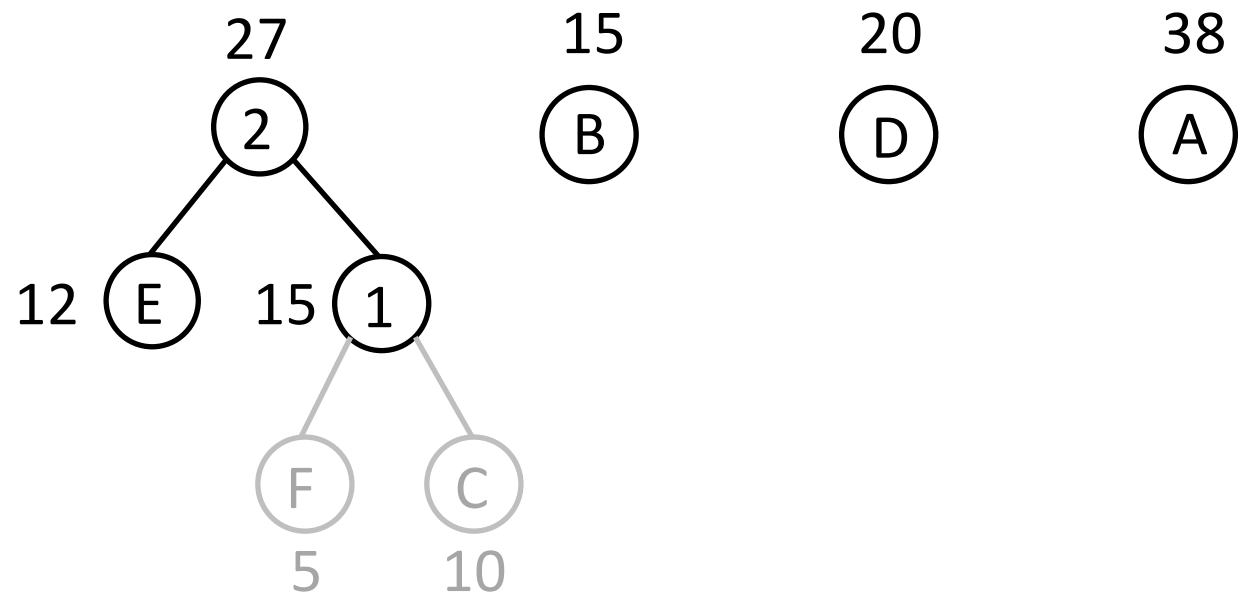
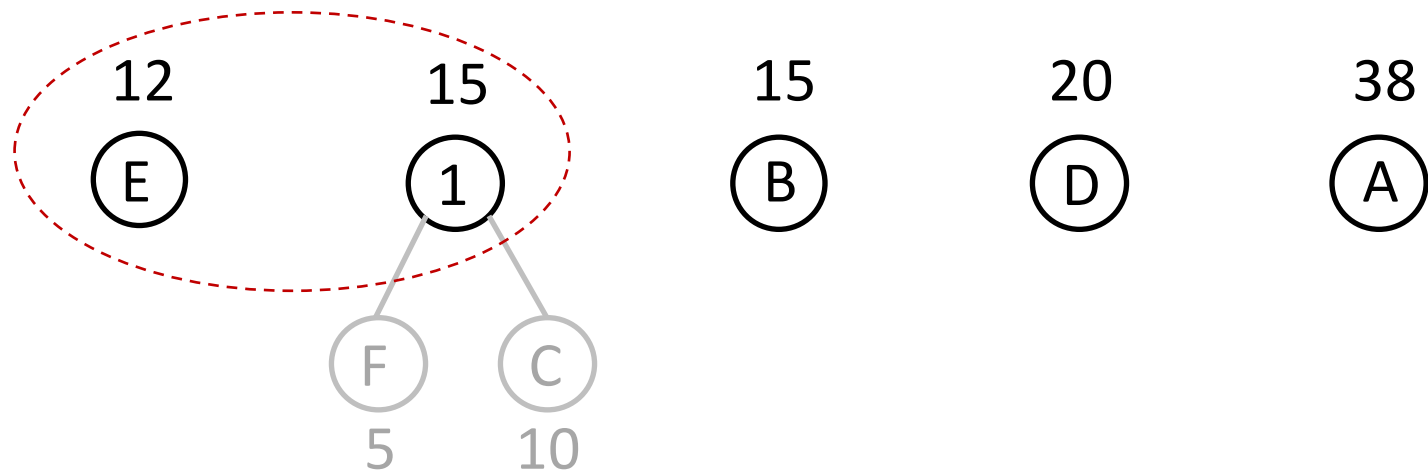


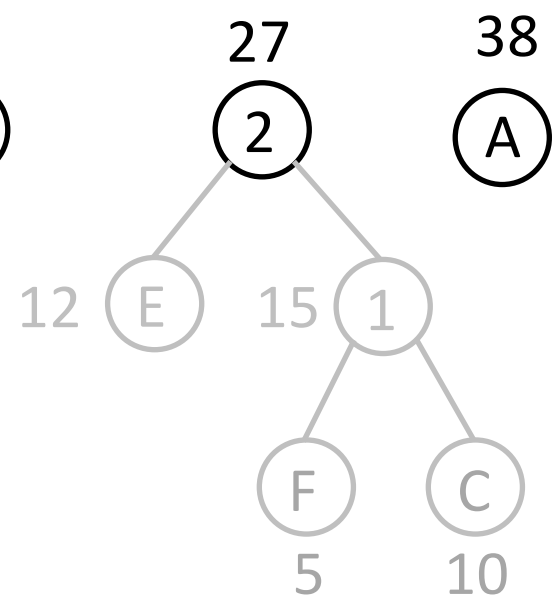
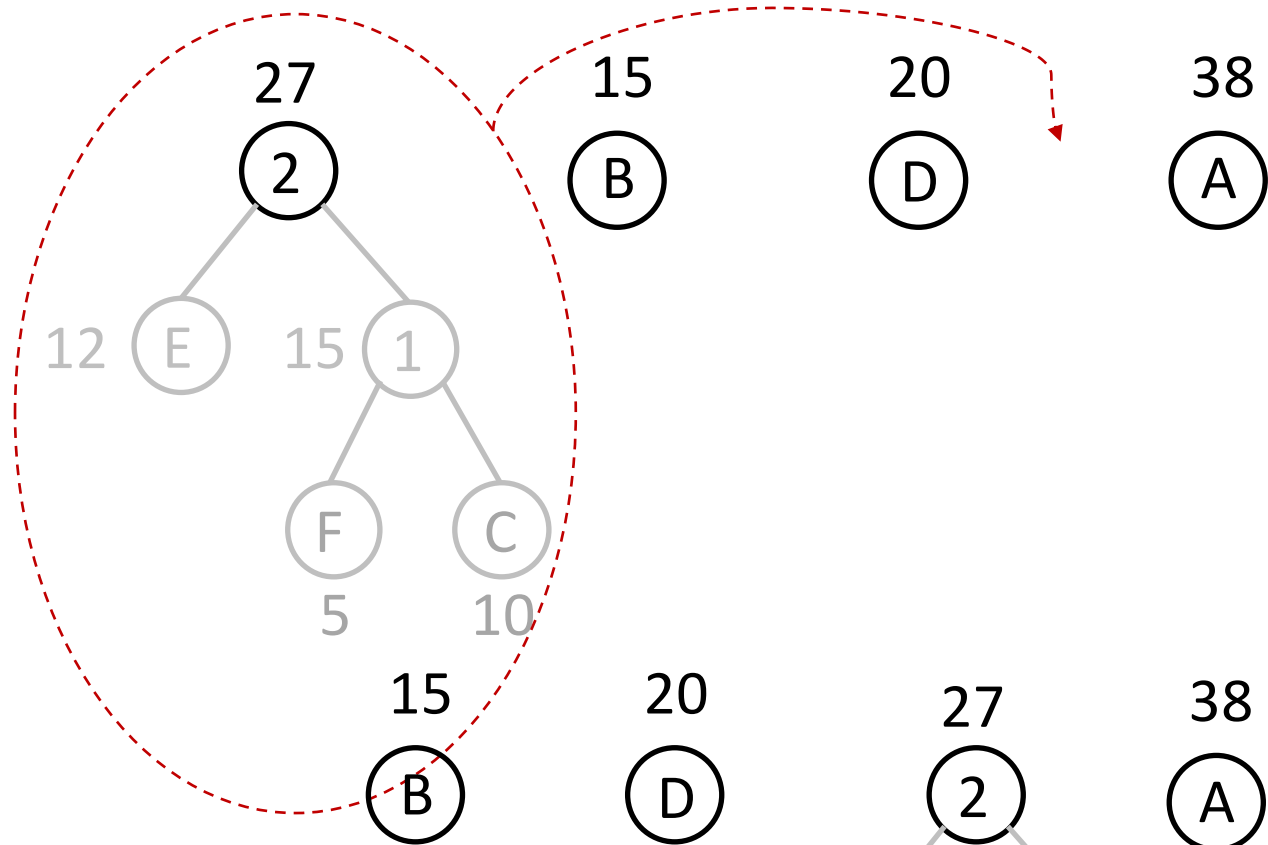
Symbol	Frequency	Codeword
A	0.38	1
B	0.15	011
C	0.10	0000
D	0.20	010
E	0.12	001
F	0.05	0001

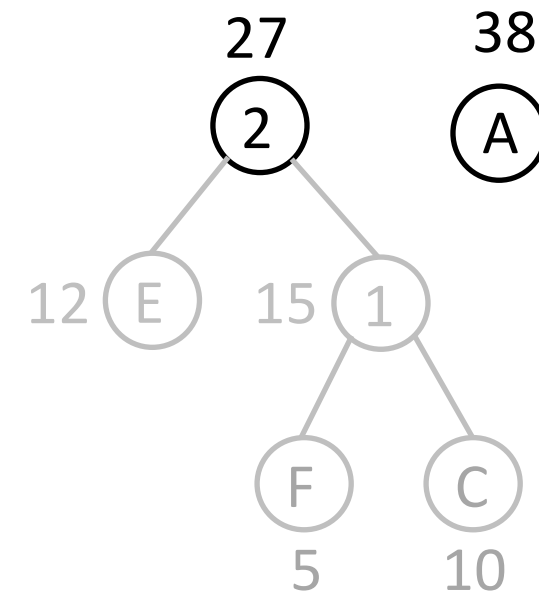
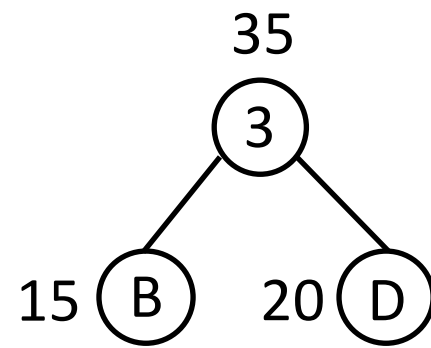
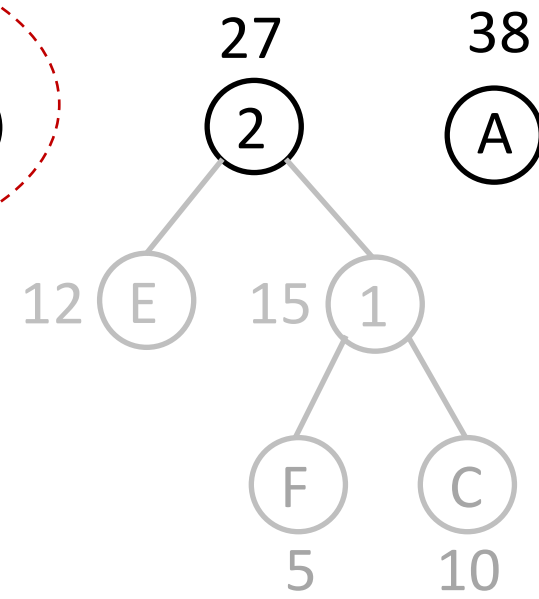
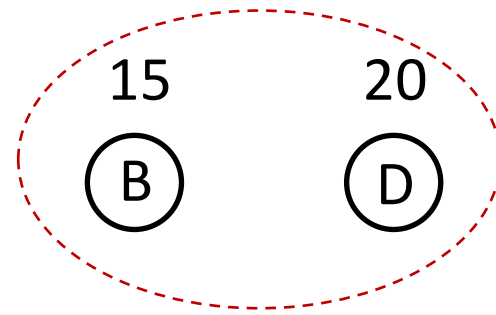
$$AD(T_3) = 0.10 \times 4 + 0.05 \times 4 + 0.12 \times 3 + 0.20 \times 3 + 0.15 \times 3 + 0.38 \times 1 = 2.39$$

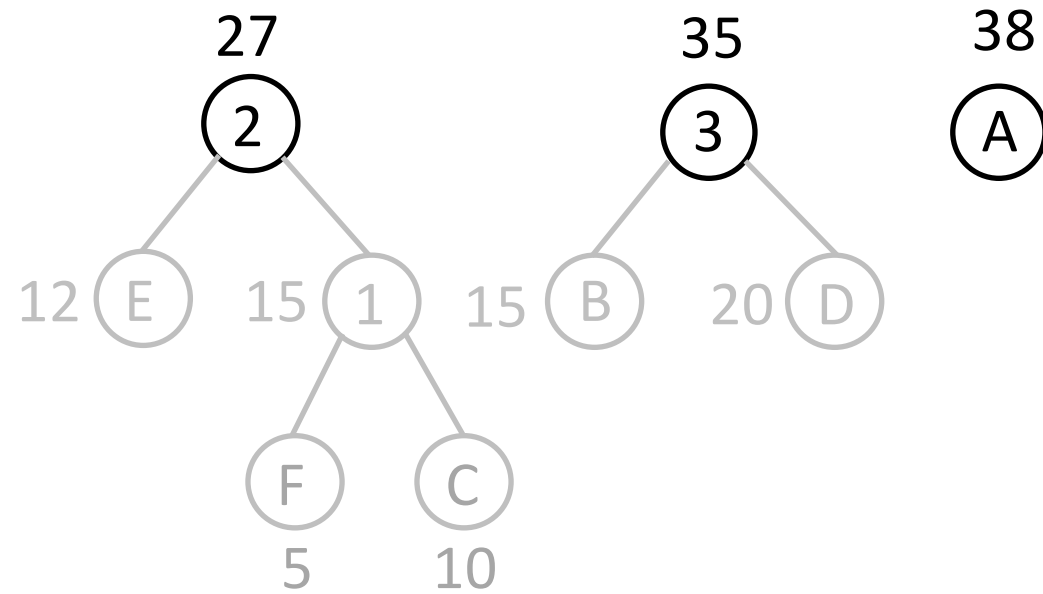
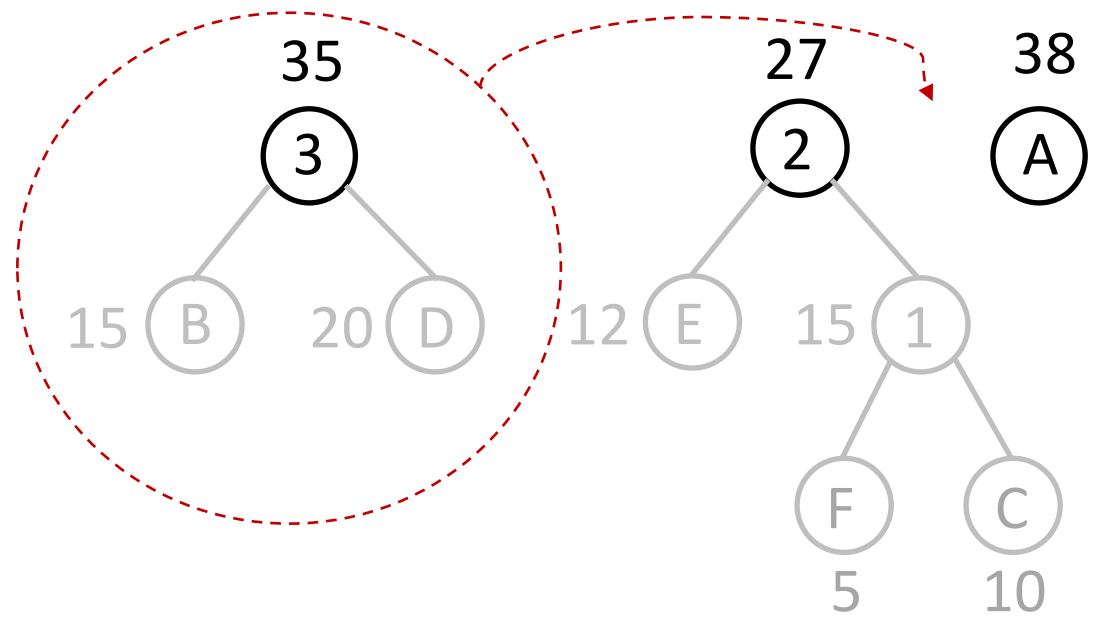


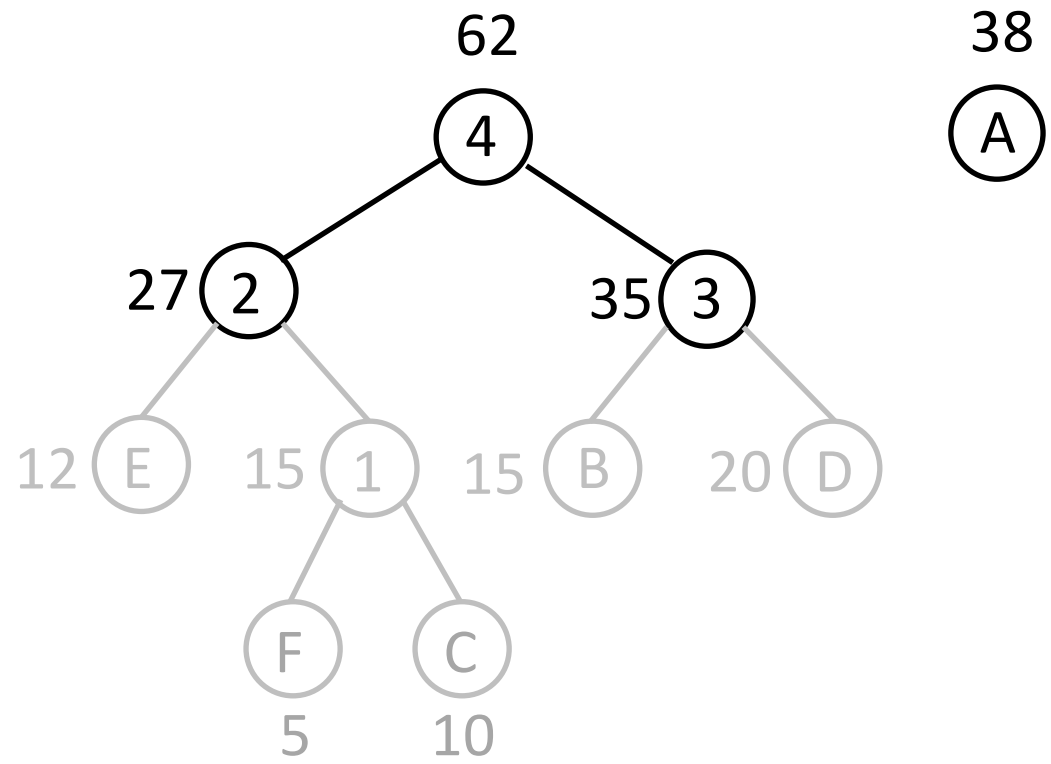
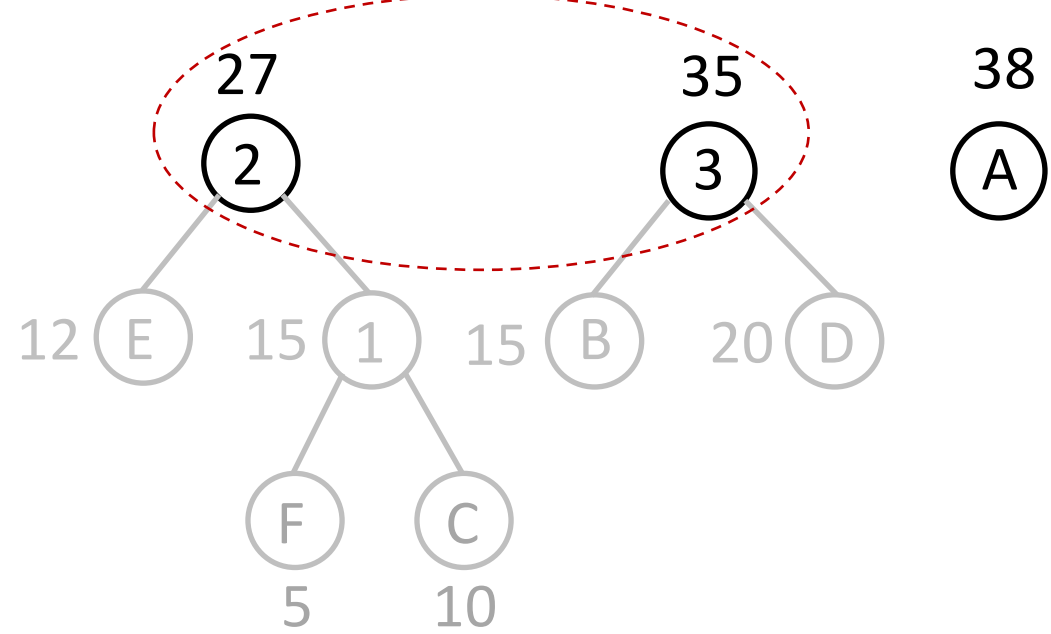


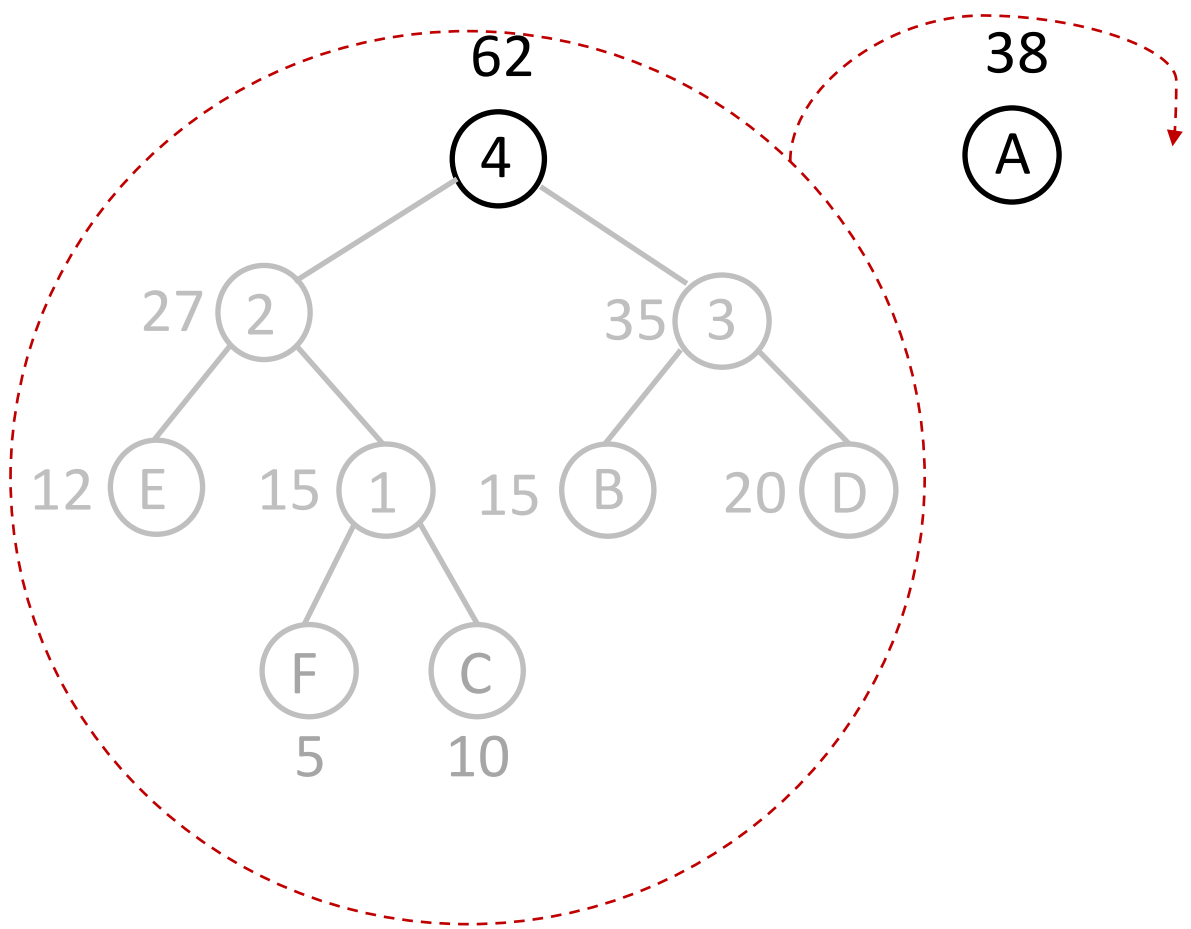




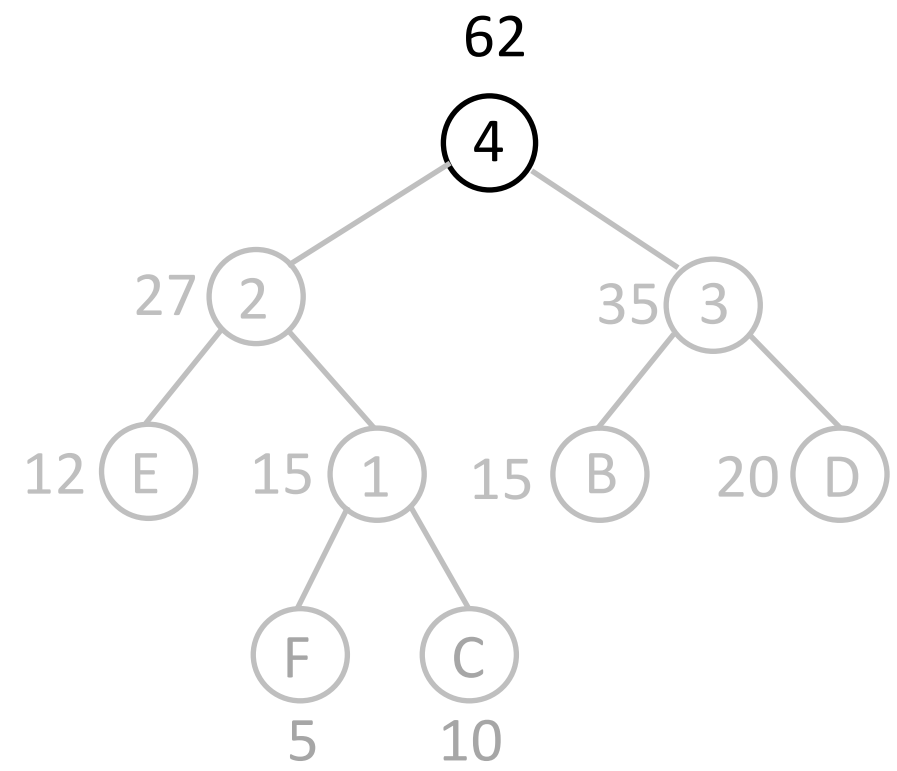


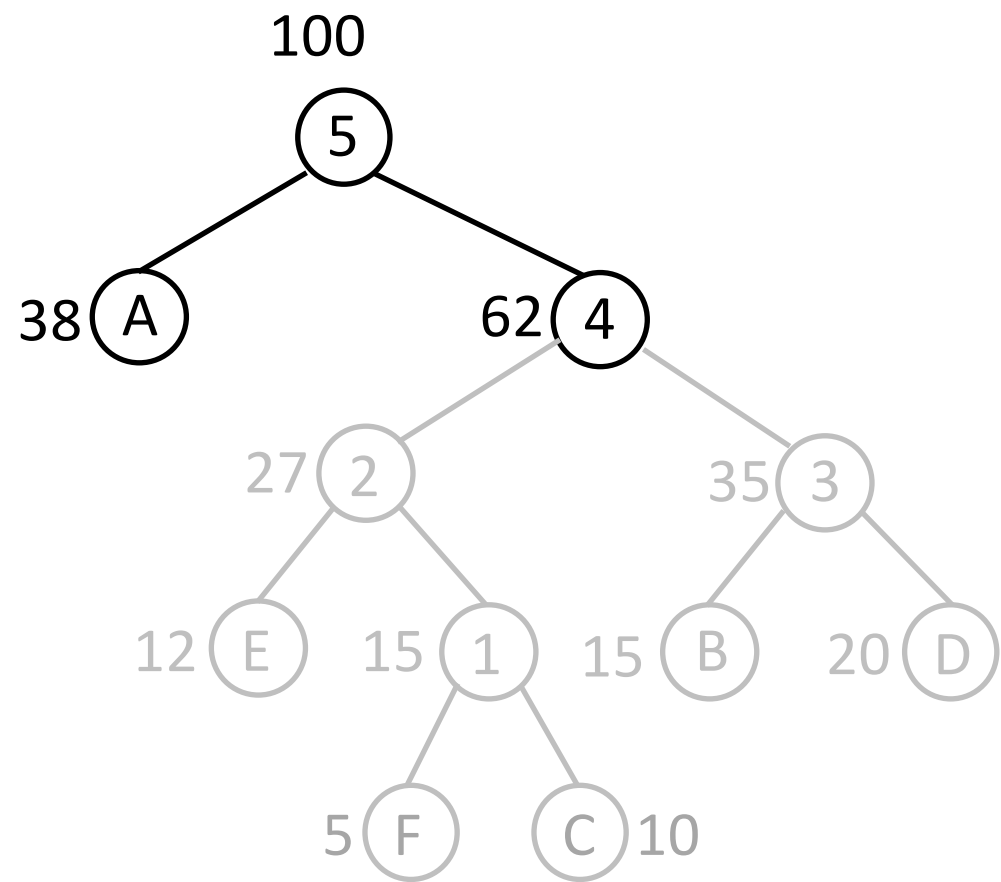
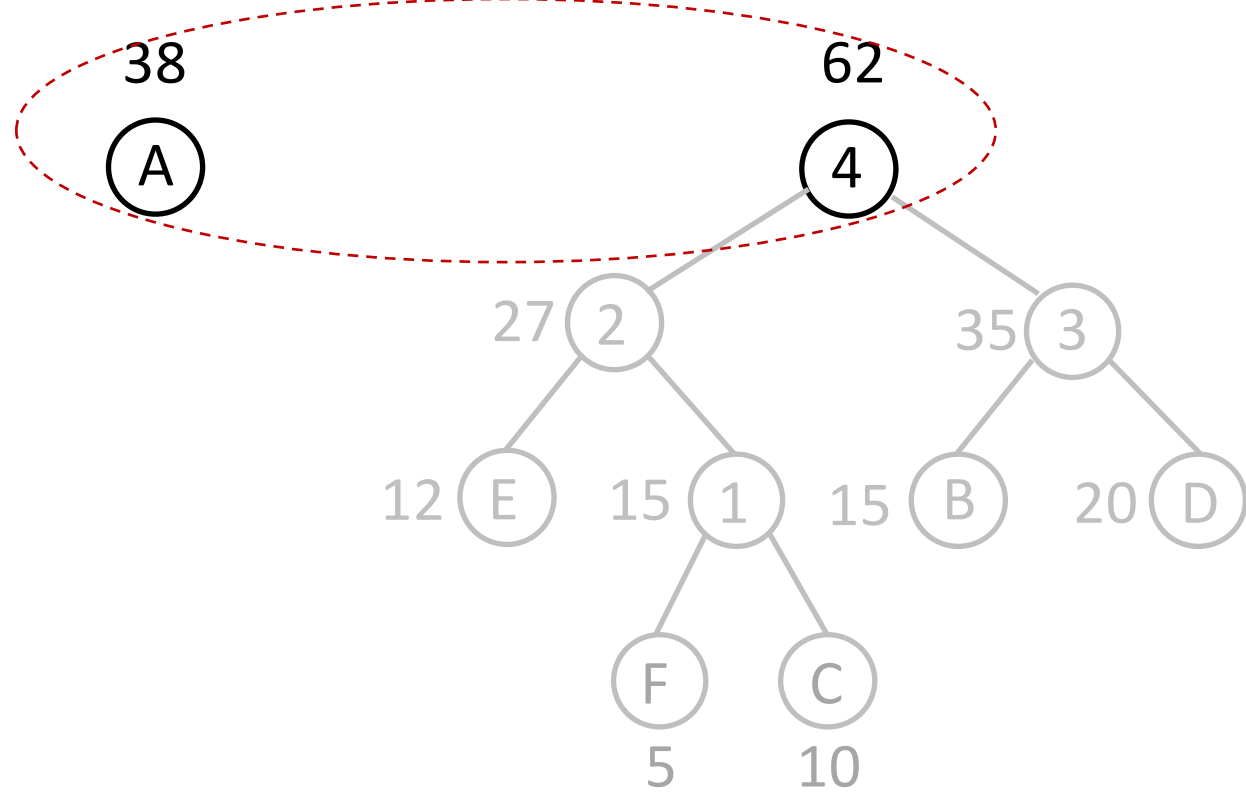


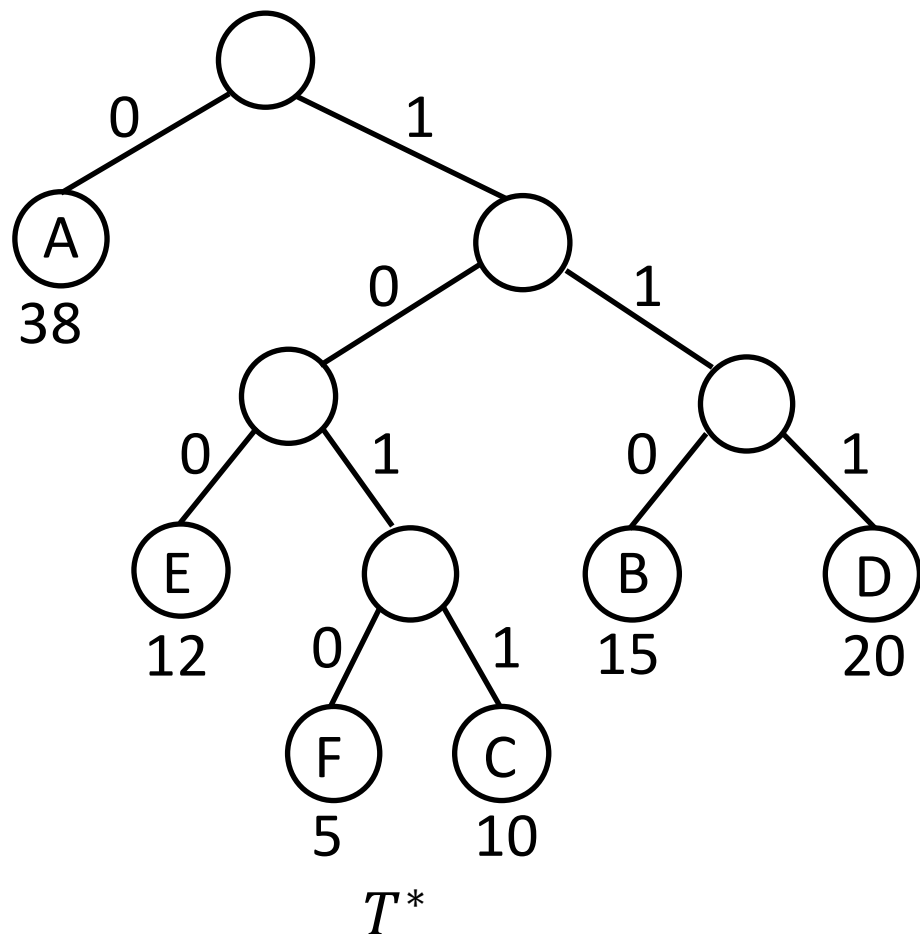




38
A







Symbol	Frequency	Codeword
A	.38	0
B	.15	110
C	.10	1011
D	.20	111
E	.12	100
F	.05	1010

$$AT(T^*) = 0.10 \times 4 + 0.05 \times 4 + 0.12 \times 3 + 0.20 \times 3 + 0.15 \times 3 + 0.38 \times 1 = 2.39$$