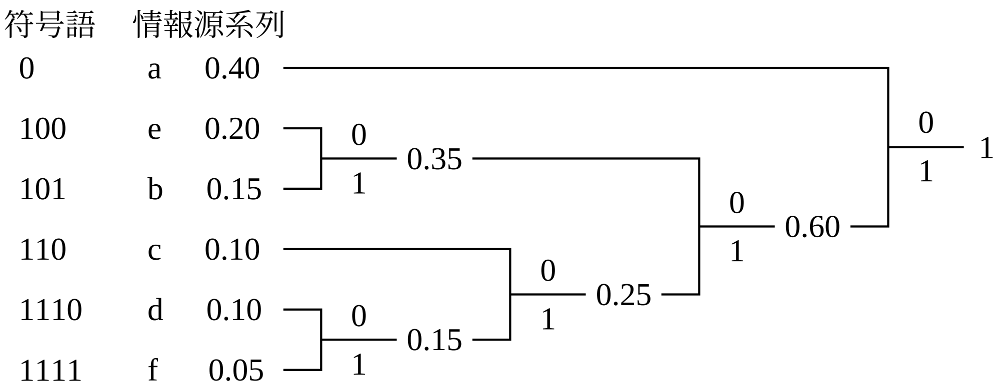
情報通信システム 小テスト第 1 回 解答

【問 1】情報源記号 a, b, c, d, e, f が生起する定常無記憶情報源 S に関して，以下の問いに答えよ．（16 点）

(1) a, b, c, d, e, f の生起確率がそれぞれ $0.40, 0.15, 0.10, 0.10, 0.20, 0.05$ のとき，2 元ハフマン符号化をせよ．また，求めた符号の 1 情報源記号あたりの平均符号長を求めよ．

（解答） たとえば，以下のようにハフマン符号化できる．



よって，1 情報源記号あたりの平均符号長 L は以下のように計算できる．

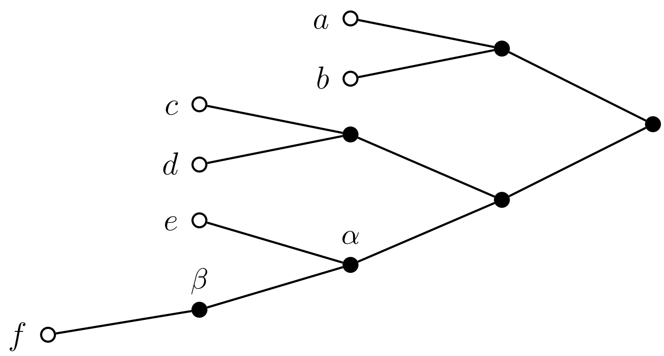
$$L = 0.4 \times 1 + (0.2 + 0.15 + 0.1) \times 3 + (0.1 + 0.05) \times 4 = 2.35$$

(2) 情報源 S の 2 元符号で各符号長が $2, 2, 3, 3, 3, 4$ となる瞬時符号は存在するか？理由とともに答えよ．

（解答） $2^{-2} + 2^{-2} + 2^{-3} + 2^{-3} + 2^{-3} + 2^{-4} = 1/2 + 3/8 + 1/16 = 15/16 < 1$ より，クラフトの不等式を満たすので，このような 2 元瞬時符号は存在する．

(3) 情報源 S を 2 元ハフマン符号化したときに，各符号長が $2, 2, 3, 3, 3, 4$ となるような a, b, c, d, e, f の生起確率分布は存在するか？存在する場合には，例を一つ示せ．存在しない場合には，その理由を答えよ．

（解答） 各符号長が $2, 2, 3, 3, 3, 4$ となる符号の木は以下のようになる．



これは明らかにハフマン符号化によって得られるコンパクト符号の木ではない．なぜならば，長さ 4 の符号を割り当てられている f に長さ 3 の符号を割り当てる（節点 β を削除して，節点 α の子とする）ことによって，平均符号長が短くなるからである．したがって，このような符号長のハフマン符号を生成する生起確率分布は存在しない．

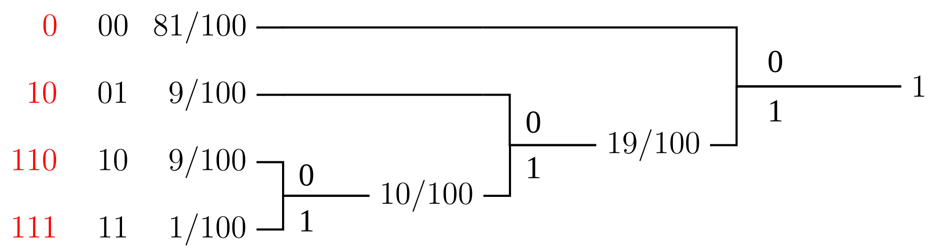
【問 2】0, 1 がそれぞれ 0.9, 0.1 の確率で生起する定常無記憶情報源 S について，以下の問いに答えよ.（34 点）

(1) 情報源 S のエントロピー $H(S)$ を求めよ.

（解答） $H(S) = -1/10 \log(1/10) - 9/10 \log(9/10) = \log 10 - 9/10 \log 9 = 1 + \log 5 - \dfrac{9}{5} \log 3$

(2) 情報源 S の 2 次の拡大情報源 S^2 を 2 元ハフマン符号化したときの 1 情報源記号あたりの平均符号長を求めよ.

（解答） 以下の図より，平均符号長 $L* = 81/100 \times 1 + 9/100 \times 2 + 10/100 \times 3 = 129/100$ なので，1 情報源記号あたりの平均符号長は，$L = L*/2 = 129/200 (= 0.645)$ となる.



(3) 3 次の拡大情報源 S^3 のエントロピー $H(S^3)$ を求めよ.

（解答） 補助定理 2.4 より $H(S^3) = 3H(S)$ が成立するので，

$$H(S^3) = 3 + 3 \log 5 - \frac{27}{5} \log 3$$

(4) n を次第に大きくしていったとき，n 次の拡大情報源 S^n の 1 情報源記号あたりの平均符号長はある値に近づいていく．その値を理由とともに示せ.

（解答） 情報源符号化定理より，n を大きくしていけば限りなく情報源 S のエントロピー $H(S) = 1 + \log 5 - 9/5 \log 3$ に近づく.

(5) 情報源 S に対して長さ 4 までの 0 のランをハフマン符号化したときの 1 情報源記号あたりの平均符号長を求めよ．

（解答） 下図より，平均符号長は，$L* = 6561/10^4 \times 1 + 3439/10^4 \times 3 = 16878/10000$ となる．一方，平均系列長は，$\overline{n} = 1000/10^4 \times 1 + 900/10^4 \times 2 + 810/10^4 \times 3 + (6561 + 729)/10^4 \times 4 = 34390/10000$ となる．したがって，1 情報源記号あたりの平均符号長は，$L = L * /\overline{n} = 16878/34390 = 8439/17195(\simeq 0.491)$ となる．