Solutions to MATH3411 Problems 35-42

35. Here, it is a good idea to draw the decision tree arising from the Huffman algorithm. However, I am going to be lazy and just write up the steps without drawing anything.
Let us first find the binary Huffman code for $S^1 = S = \{s_1, s_2\}$:

Source	Step 0	Step 1
s_1	$p_1 = \frac{2}{3}$ **0**	$p_{12} = 1$ $\emptyset$
s_2	$p_2 = \frac{1}{3}$ **1**	

The binary Huffman code for S^1 is 0, 1 . The expected codeword length is 1.
Let us now find the binary Huffman code for $S^2 = \{s_1s_1, s_1s_2, s_2s_1, s_2s_2\}$:

Source	Step 0	Step 1	Step 2	Step 3
$\sigma_1 = s_1s_1$	$p_1 = \frac{4}{9}$ **1**	$p_1 = \frac{4}{9}$ **1**	$p_{342} = \frac{5}{9}$ **0**	$p_{3421} = 1$ $\emptyset$
$\sigma_2 = s_1s_2$	$p_2 = \frac{2}{9}$ **01**	$p_{34} = \frac{3}{9}$ **00**	$p_1 = \frac{4}{9}$ **1**	
$\sigma_3 = s_2s_1$	$p_3 = \frac{2}{9}$ **000**	$p_2 = \frac{2}{9}$ **01**		
$\sigma_4 = s_2s_2$	$p_4 = \frac{1}{9}$ **001**			

In other words, the binary Huffman code for $\sigma_1, \ldots, \sigma_4$ is 1, 01, 000, 001 .
The expected codeword length is $1 + \frac{5}{9} + \frac{3}{9} = \frac{17}{9} \approx 1.889$ (using Knuth's theorem).
Finally, let us find the binary Huffman code for $S^3 = \{s_1s_1s_1, s_1s_1s_2, s_1s_2s_1, s_2s_1s_1, s_1s_2s_2, s_2s_1s_2, s_2s_2s_1, s_1s_1s_1\}$:

Source	Step 0	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7
$\sigma_1 = s_1s_1s_1$	$p_1 = \frac{8}{27}$ **01**	$p_1 = \frac{8}{27}$ **01**	$p_1 = \frac{8}{27}$ **01**	$p_1 = \frac{8}{27}$ **01**	$p_{23} = \frac{8}{27}$ **00**	$p_{47856} = \frac{11}{27}$ **1**	$p_{231} = \frac{16}{27}$ **0**	$p_{23147856} = 1$ $\emptyset$
$\sigma_2 = s_1s_1s_2$	$p_2 = \frac{4}{27}$ **000**	$p_2 = \frac{4}{27}$ **000**	$p_{56} = \frac{4}{27}$ **11**	$p_{478} = \frac{7}{27}$ **10**	$p_1 = \frac{8}{27}$ **01**	$p_{23} = \frac{8}{27}$ **00**	$p_{47856} = \frac{11}{27}$ **1**	
$\sigma_3 = s_1s_2s_1$	$p_3 = \frac{4}{27}$ **001**	$p_3 = \frac{4}{27}$ **001**	$p_2 = \frac{4}{27}$ **000**	$p_{56} = \frac{4}{27}$ **11**	$p_{478} = \frac{7}{27}$ **10**	$p_1 = \frac{8}{27}$ **01**		
$\sigma_4 = s_2s_1s_1$	$p_4 = \frac{4}{27}$ **100**	$p_4 = \frac{4}{27}$ **100**	$p_3 = \frac{4}{27}$ **001**	$p_2 = \frac{4}{27}$ **000**	$p_{56} = \frac{4}{27}$ **11**			
$\sigma_5 = s_1s_2s_2$	$p_5 = \frac{2}{27}$ **110**	$p_{78} = \frac{3}{27}$ **101**	$p_4 = \frac{4}{27}$ **100**	$p_3 = \frac{4}{27}$ **001**				
$\sigma_6 = s_2s_1s_2$	$p_6 = \frac{2}{27}$ **111**	$p_5 = \frac{2}{27}$ **110**	$p_{78} = \frac{3}{27}$ **101**					
$\sigma_7 = s_2s_2s_1$	$p_7 = \frac{2}{27}$ **1010**	$p_6 = \frac{2}{27}$ **111**						
$\sigma_8 = s_2s_2s_2$	$p_8 = \frac{1}{27}$ **1011**							

In other words, the binary Huffman code for $\sigma_1, \ldots, \sigma_8$ is

01, 000, 001, 100, 110, 111, 1010, 1011

By Knuth's theorem, the expected codeword length then:

$$L = 1 + \frac{16}{27} + \frac{11}{27} + \frac{8}{27} + \frac{7}{27} + \frac{4}{27} + \frac{3}{27} = \frac{76}{27} \approx 2.81$$

The average codeword length per binary symbol for S^1, S^2, and S^3 is

$$L^{(1)} = 1 \qquad \frac{L^{(2)}}{2} = \frac{17}{18} \approx 0.944 \qquad \frac{L^{(3)}}{3} = \frac{76}{81} \approx 0.938$$

36.

a) The characteristic polynomial of M is

$$p_M(x) = \det(M - xI) = \det \begin{pmatrix} \frac{1}{3} - x & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{2} - x & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{2} - x \end{pmatrix} = -x^3 + \frac{4}{3}x^2 - \frac{17}{48}x + \frac{1}{48} = \frac{1}{48}(x-1)(12x-1)(4x-1)$$

We see that M has eigenvalues 1, $\frac{1}{4}$, and $\frac{1}{12}$.
Let us now solve $(M - I)\mathbf{p} = 0$ to find the equilibrium probabilities $\mathbf{p}$:

$$\begin{pmatrix} -\frac{2}{3} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{3} & -\frac{1}{2} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & -\frac{1}{2} \end{pmatrix} \xrightarrow[R_3 = R_1 + \frac{1}{2}R_1]{R_2 = R_2 + \frac{1}{2}R_1} \begin{pmatrix} -\frac{2}{3} & \frac{1}{4} & \frac{1}{4} \\ 0 & -\frac{3}{8} & \frac{3}{8} \\ 0 & \frac{3}{8} & -\frac{3}{8} \end{pmatrix} \xrightarrow[R1 = -\frac{3}{2}R1]{R_3 = R_3 + R_2} \begin{pmatrix} 1 & -\frac{3}{8} & -\frac{3}{8} \\ 0 & -\frac{3}{8} & \frac{3}{8} \\ 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow[R2 = -\frac{8}{3}]{R_1 = R_1 - R_2} \begin{pmatrix} 1 & 0 & -\frac{3}{4} \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

The eigenvectors of M for the eigenvalue 1 are thus $t\begin{pmatrix} \frac{3}{4} \\ 1 \\ 1 \end{pmatrix}$ for all $t \neq 0$.
Of these, $\mathbf{p}$ is the one with $1 = t\frac{3}{4} + t + t = t\frac{11}{4}$; i.e., $t = \frac{4}{11}$.
Therefore, $\mathbf{p} = \frac{1}{11}\begin{pmatrix} 3 \\ 4 \\ 4 \end{pmatrix}$.

b) There might well be a better way to answer this (without referring to Markov literature) but here is one way. Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be eigenvectors of M for $\frac{1}{4}$ and $\frac{1}{2}$, respectively. Then we can diagonalise M as $M = N^{-1}DM$ where N has columns $\mathbf{p}, \mathbf{v}_1, \mathbf{v}_2$, and $D = \operatorname{diag}(1, \frac{1}{4}, \frac{1}{2})$. Then

$$\lim_{n\to\infty} M^n = \lim_{n\to\infty} (N^{-1}DN)^n = \lim_{n\to\infty} N^{-1}D^nN = \lim_{n\to\infty} N^{-1} \begin{pmatrix} 1^n & 0 & 0 \\ 0 & \left(\frac{1}{4}\right)^n & 0 \\ 0 & 0 & \left(\frac{1}{2}\right)^n \end{pmatrix} N = N^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} N$$

We see that $M^\infty := \lim n \to \infty M^n$ does indeed exist. To see what M^∞ is, you might be able to stare hard at the expression above; to be honest, I can't see it myself but it probably isn't difficult. Instead, let us calculate M^∞ in another way: First note that $M^\infty\mathbf{p} = \lim_{n\to\infty} M^n\mathbf{p} = \lim_{n\to\infty} p = p$, so M^∞ has $\mathbf{p}$ as eigenvector for the eigenvalue 1. Similarly, $\mathbf{v}_1$ and $\mathbf{v}_2$ are eigenvectors of M^∞ with eigenvalues $\frac{1}{4}$ and $\frac{1}{2}$. Thus, M^∞ has three eigenvalues, the same number as the size of M, so the dimension of the eigenspace E_1 for the eigenvalue 1 is 1. In other words, the eigenvectors of M^∞ for the eigenvalue 1 is the span of $\mathbf{p}$. There is thus only the vector $\mathbf{p}$ that is an eigenvector for 1 and has unit length. Now note that $MM^\infty = M^\infty$. Letting A_i denote the ith column of any matrix A, we see that, for any $i = 1, 2, 3$, $M(M^\infty)_i = (MM^\infty)_i = (M^\infty)_i$. In other words, each of the three columns $(M^\infty)_i$ is an eigenvector for the eigenvalue 1; they must therefore be scalar multiples of $\mathbf{p}$. However, each column of M^∞ has unit length (since the columns of M have unit length, M^T has the all-1 vector as eigenvector, as does therefore M^∞) and must therefore equal $\mathbf{p}$. To conclude, the columns of M^∞ all equal $\mathbf{p}$.

37.

a) Let us now solve $(M-I)\mathbf{p}=0$ to find the equilibrium vector $\mathbf{p}$:

$$\begin{pmatrix} -0.3 & 0.2 & 0.1 \\ 0.2 & -0.4 & 0.4 \\ 0.1 & 0.2 & -0.5 \end{pmatrix} \xrightarrow[R_2=R_2-2R_3]{R_1=R_1+3R_3} \begin{pmatrix} 0 & 0.8 & -1.4 \\ 0 & -0.8 & 1.4 \\ 0.1 & 0.2 & -0.5 \end{pmatrix} \xrightarrow[R3=4R3]{R1=R_1+R_2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & -0.8 & 1.4 \\ 0.4 & 0.8 & -2 \end{pmatrix}$$

$$\xrightarrow{R_3=R_3+R_2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & -0.8 & 1.4 \\ 0.4 & 0 & -0.6 \end{pmatrix}$$

The eigenvectors of M for the eigenvalue 1 are thus $t(\frac{0.6}{0.4}, \frac{1.4}{0.8}, 1)^T = t(\frac{3}{2}, \frac{7}{4}, 1)^T$ for all $t \neq 0$.
Of these, $\mathbf{p}$ is the one with $1 = t\frac{7}{4} + \frac{3}{2}t + t = t\frac{17}{4}$; i.e., $t = \frac{4}{17}$.

Therefore, $\mathbf{p} = \frac{1}{17}\begin{pmatrix} 6 \\ 7 \\ 4 \end{pmatrix}$.

Let us now calculate the Huffman codes Huff_{E}, $\text{Huff}_{(2)}$, $\text{Huff}_{(2)}$, $\text{Huff}_{(3)}$:

Source	p_i	Huff_E
s_1	$\frac{6}{17}$	00
s_2	$\frac{7}{17}$	1
s_3	$\frac{4}{17}$	01

Source	p_i	$\text{Huff}_{(1)}$
s_1	0.7	0
s_2	0.2	10
s_3	0.1	11

Source	p_i	$\text{Huff}_{(2)}$
s_1	0.2	10
s_2	0.6	0
s_3	0.2	11

Source	p_i	$\text{Huff}_{(3)}$
s_1	0.1	01
s_2	0.4	00
s_3	0.5	1

The average lengths of these codes

$$L_E = \frac{28}{17} \approx 1.65 \qquad L_{(1)} = 1.3 \qquad L_{(2)} = 1.4 \qquad L_{(3)} = 1.6$$

The Markov Huffman code has average length

$$L_M = \frac{6}{17}L_{(1)} + \frac{7}{17}L_{(2)} + \frac{4}{17}L_{(3)} = \frac{6}{17}1.3 + \frac{7}{17}1.4 + \frac{4}{17}1.6 \approx 1.41$$

This is less than $L_E \approx 1.65$, and only about $\frac{L_M}{L_{\text{block}}} = \frac{1.41}{2} \approx 71\%$ of the block code length.

b) We encode $s_2s_2s_1s_1s_2s_3s_3$:

symbol	code to use	encoded symbol
s_2	Huff_E	1
s_2	$\text{Huff}_{(2)}$	0
s_1	$\text{Huff}_{(2)}$	10
s_1	$\text{Huff}_{(1)}$	0
s_2	$\text{Huff}_{(1)}$	10
s_3	$\text{Huff}_{(2)}$	11
s_3	$\text{Huff}_{(3)}$	1

so this is encoded as 1010010111 .

c) We decode 010001010 :

code to use	encoded symbol	decoded symbol
Huff_E	01	s_3
$\text{Huff}_{(3)}$	00	s_2
$\text{Huff}_{(2)}$	0	s_2
$\text{Huff}_{(2)}$	10	s_1
$\text{Huff}_{(1)}$	10	s_2

so this is decoded as $s_3 s_2 s_2 s_1 s_2$.

38. We wish to encode the message $bac\bullet$:

	subinterval start	**width**
symbols	0	1
b	$0 + \frac{2}{5} \times 1 = \frac{2}{5}$	$\frac{1}{5} \times 1 = \frac{1}{5}$
a	$\frac{2}{5} + 0 \times \frac{1}{5} = \frac{2}{5}$	$\frac{2}{5} \times \frac{1}{5} = \frac{2}{25}$
c	$\frac{2}{5} + \frac{3}{5} \times \frac{2}{25} = \frac{56}{125}$	$\frac{1}{5} \times \frac{2}{25} = \frac{2}{125}$
$\bullet$	$\frac{56}{125} + \frac{4}{5} \times \frac{2}{125} = \frac{288}{625}$	$\frac{1}{5} \times \frac{2}{125} = \frac{2}{625}$

We must therefore choose a number in the interval $[\frac{288}{625}, \frac{290}{625}) = [0.4608, 0.4640)$, like 0.4620 say.

39.

a) We wish to encode the message $s_2 s_1 s_3 s_1 \bullet$:

	subinterval start			**width**
begin			0	1
s_2	$0 + 0.4 \times 1$	$=$	0.4	$0.3 \times 1 = 0.3$
s_1	$0.4 + 0 \times 0.3$	$=$	0.4	$0.4 \times 0.3 = 0.12$
s_3	$0.4 + 0.7 \times 0.12$	$=$	0.484	$0.2 \times 0.12 = 0.024$
s_1	$0.484 + 0 \times 0.024$	$=$	0.484	$0.4 \times 0.024 = 0.0096$
$\bullet$	$0.484 + 0.9 \times 0.0096$	$=$	0.49264	$0.1 \times 0.096 = 0.00096$

We must pick a number in the interval $[0.49264, 0.49264 + 0.00096) = [0.49264, 0.4936)$, like 0.493 say.

b) We wish to decode the number 0.12345 :

code number rescaled	**in interval**	**decoded symbol**
0.12345	$[0, 0.4)$	s_1
$(0.12345 - 0)/.4 = 0.308625$	$[0, 0.4)$	s_1
$(0.308625 - 0)/.4 = 0.7715625$	$[0.7, 0.9)$	s_3
$(0.7715625 - 0.7)/.2 = 0.3578125$	$[0, 0.4)$	s_1
$(0.3578125 - 0)/.4 = 0.89453125$	$[0.7, 0.9)$	s_3
$(0.89453125 - 0.7)/.4 = 0.97265625$	$[0.9, 1)$	$\bullet$

The decoded message is then $s_1 s_1 s_3 s_1 s_3 \bullet$.

40.

a) We wish to encode the message ma␣na␣ma␣na␣doo␣doo␣doo␣doo␣doo␣ma␣na␣ma␣na␣doo␣doo␣doo␣ :

r	s	ℓ	**new entry**	**output**
ma␣na␣ma␣na␣doo␣doo␣doo␣doo␣doo␣ma␣na␣ma␣na␣doo␣doo␣doo␣doo	∅	0	1. m	(0,m)
a␣na␣ma␣na␣doo␣doo␣doo␣doo␣doo␣ma␣na␣ma␣na␣doo␣doo␣doo␣doo	∅	0	2. a	(0,a)
␣na␣ma␣na␣doo␣doo␣doo␣doo␣doo␣ma␣na␣ma␣na␣doo␣doo␣doo␣doo	∅	0	3. ␣	(0,␣)
na␣ma␣na␣doo␣doo␣doo␣doo␣doo␣ma␣na␣ma␣na␣doo␣doo␣doo␣doo	∅	0	4. n	(0,n)
a␣ma␣na␣doo␣doo␣doo␣doo␣doo␣ma␣na␣ma␣na␣doo␣doo␣doo␣doo	a	2	5. a␣	(2,␣)
ma␣na␣doo␣doo␣doo␣doo␣doo␣ma␣na␣ma␣na␣doo␣doo␣doo␣doo	m	1	6. ma	(1,a)
␣na␣doo␣doo␣doo␣doo␣doo␣ma␣na␣ma␣na␣doo␣doo␣doo␣doo	␣	3	7. ␣n	(3,n)
a␣doo␣doo␣doo␣doo␣doo␣ma␣na␣ma␣na␣doo␣doo␣doo␣doo	a␣	5	8. a␣d	(5,d)
oo␣doo␣doo␣doo␣doo␣ma␣na␣ma␣na␣doo␣doo␣doo␣doo	∅	0	9. o	(0,o)
o␣doo␣doo␣doo␣doo␣ma␣na␣ma␣na␣doo␣doo␣doo␣doo	o	9	10. o␣	(9,␣)
doo␣doo␣doo␣doo␣ma␣na␣ma␣na␣doo␣doo␣doo␣doo	∅	0	11. d	(0,d)
oo␣doo␣doo␣doo␣ma␣na␣ma␣na␣doo␣doo␣doo␣doo	o	9	12. oo	(9,o)
␣doo␣doo␣doo␣ma␣na␣ma␣na␣doo␣doo␣doo␣doo	␣	3	13. ␣d	(3,d)
oo␣doo␣doo␣ma␣na␣ma␣na␣doo␣doo␣doo␣doo	oo	12	14. oo␣	(12,␣)
doo␣doo␣ma␣na␣ma␣na␣doo␣doo␣doo␣doo	do	11	15. do	(11,o)
o␣doo␣ma␣na␣ma␣na␣doo␣doo␣doo␣doo	o␣	10	16. o␣d	(10,d)
oo␣ma␣na␣ma␣na␣doo␣doo␣doo␣doo	oo␣	14	17. oo␣m	(14,m)
a␣na␣ma␣na␣doo␣doo␣doo␣doo	a␣	5	18. a␣n	(5,n)
a␣ma␣na␣doo␣doo␣doo␣doo	a␣	5	19. a␣m	(5,m)
a␣na␣doo␣doo␣doo␣doo	a␣n	18	20. a␣na	(18,a)
␣doo␣doo␣doo␣doo	␣d	13	21. ␣do	(13,o)
o␣doo␣doo␣doo	o␣d	16	22. o␣do	(16,o)
o␣doo␣doo	o␣do	22	23. o␣doo	(22,o)
␣doo	␣do	21	24. ␣doo	(21,o)

The encoded message is then

(0,m)(0, a)(0,␣)(0, n)(2,␣)(1, a)(3, n)(5, d)(0, o), (9,␣)(0, d)(9, o)(3, d)(12,␣)(11, o)(10, d)(14,m)(5, n)(5,m)(18, a)(13, o)(16, o)(22, o)(21, o)

b) Let us decode the codeword (0,t)(0,o)(0,␣)(0,b)(0,e)(3,o)(0,r)(3,n)(2,t)(3,t)(2,␣)(4,e) :

output	**new dictionary entry**
(0,t)	1. t
(0,o)	2. o
(0,␣)	3. ␣
(0,b)	4. b
(0,e)	5. e
(3,o)	6. ␣o
(0,r)	7. r
(3,n)	8. ␣n
(2,t)	9. ot
(3,t)	10. ␣t
(2,␣)	11. o␣
(4,e)	12. be

The decoded text is “to be or not to be”.

41. $H(S) = -\frac{2}{3}\log_2\frac{1}{3} - \frac{3}{9}\log_2\frac{1}{9} = \frac{2}{3}\log_2 3 + \frac{6}{9}\log_2 3 = \frac{4}{3}\log_2 3 \approx 2.113$

42.

Q29: **a)** $H(S) = -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{3}\log_2\frac{1}{3} - \frac{1}{6}\log_2\frac{1}{6} \approx 1.459$
The average length $L = 1.5 > H(S)$ - but pretty close.

b) $H(S) = -\frac{1}{3}\log_2\frac{1}{3} - \frac{1}{4}\log_2\frac{1}{4} - \frac{1}{5}\log_2\frac{1}{5} - \frac{1}{6}\log_2\frac{1}{6} - \frac{1}{20}\log_2\frac{1}{20} \approx 2.140$
The average length $L = 2.217 > H(S)$ - but not far off.

c) $H(S) = -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{4}\log_2\frac{1}{4} - \frac{1}{8}\log_2\frac{1}{8} - \frac{1}{16}\log_2\frac{1}{16} - \frac{1}{16}\log_2\frac{1}{16} \approx 1.875$
The average length $L = 1.875 = H(S)$ - exactly the same!

d) $H(S) = -\frac{27}{40}\log_2\frac{27}{40} - \frac{9}{40}\log_2\frac{9}{40} - \frac{3}{40}\log_2\frac{3}{40} - \frac{1}{40}\log_2\frac{1}{40} \approx 1.280$
The average length $L = 1.425 > H(S)$ - not too far off.

Q33: $H(S) = -0.22\log_4 0.22 - 0.2\log_4 0.2 - 0.18\log_4 0.18 - 0.15\log_4 0.15 - 0.10\log_4 0.10 - 0.08\log_4 0.08 - 0.05\log_4 0.05 - 0.02\log_2 0.02 \approx 1.377$
The average length $L = 1.47 > H(S)$ - but pretty close.

Q35: $H(S^1) = -\frac{2}{3}\log_2\frac{2}{3} - \frac{1}{3}\log_2\frac{1}{3} \approx 0.918$
$H(S^2) = 2H(S^1) \approx 1.837$
$H(S^3) = 3H(S^1) \approx 2.755$
The corresponding average lengths are 1, 1.889 and 2.815, respectively.
These are all greater than the corresponding entropies - but not by a lot.

Q39: $H(S) = -0.4\log_{10} 0.4 - 0.3\log_{10} 0.3 - 0.2\log_{10} 0.2 - 0.1\log_{10} 0.1 \approx 0.5558$ digits/symbol.
The 5-symbol message $s_2 s_1 s_3 s_1 \bullet$ was encoded as 0.493, so $\frac{3}{5} = 0.6$ digits per symbol were used, which is more than 0.5558 but not by a lot.