Solutions to Assignment 1

Question 1

Many files, including PDF files, begin with a "magic" value that identifies the type of file. For this specific PDF file the magic value is the 8-byte string: "%PDF-1.5".

(a) The values of the first 8 bytes in hexadecimal are: `25 50 44 46 2d 31 2e 35`.

(b) The values of the first four 16-bit values assuming little-endian byte order are: 0x5025 (20517), 0x4644 (17988), 0x312d (12589) and 0x352e (13614).

(c) The values of the first two 32-bit values, assuming big-endian byte order are: 0x25504446 (626017350) and 0x2d312e35 (758197813).

(d) Since the first 8 characters are all less than or equal to 0x7f they are all ASCII (in the Unicode Basic Latin Table) and the characters are "%PDF-1.5".

Question 2

(a)

(b) We can compute the number of bits of information for each message as $\log_2(P)$ (or $(\log(P)/\log(2))$ and the entropy by weighting each message's information by its probability. This is easily done using a spreadsheet:

P(m)	I(m)	P(m)*I(M)
0.01	6.64	0.07
0.03	5.06	0.15
0.10	3.32	0.33
0.25	2.00	0.50
0.25	2.00	0.50
0.25	2.00	0.50
0.10	3.32	0.33
0.01	6.64	0.07
1.00	**Entropy:**	**2.45**

(c) Assuming all messages are equally likely, and one message is generated per minute, the information rate is: $2.45/60 = 41 \times 10^{-3}$ bits/second.

(d) The data rate if we transmitted 3 bits per message would be: $3/60 = 50 \times 10^{-3}$ bits/second.

Question 3

The character with Unicode code point U+1234 is Ethiopic Syllable See: ሴ.

The value 0x1234 is 0001 0010 0011 0100 in binary and must therefore be encoded into UTF-8 using the third row of Table 3-6.

The bits assignments are: zzzz=0001 yyyyyy=001000 xxxxxx=110100 and the UTF-8 encoding in binary is: 1110 0001 10 001000 10 110100 which is 0xe1 0x88 0xb4.

Question 4

The bytes 0xe5 0x93 0xb2 are 1110 0101 1001 0011 1011 0010 in binary from which we can get the bit assignments: zzzz=0101, yyyyyy=010011 and xxxxxx=110010.

The code point in binary is thus 0101 0100 1111 0010 which is 0x54f2.

From the Unicode standard this is: 哲 which means "wise, sagacious; wise man, sage."

Question 5

(a) The FER is the fraction of frames with errors or $\frac{23}{750000} \approx 31 \times 10^{-6}$.

(b) If each frame with errors had only one bit in error the number of bits is $128 \times 750,000 = 96 \times 10^6$ and the BER is the faction of bits with errors, or $\frac{23}{96 \times 10^6} \approx 240 \times 10^{-9}$.