Solutions to Assignment 1

Question 1

If 20 words have probability of 0.02 then the probability of one of these words is $20 \times 0.02 = 0.4$ thus the probability of the others is $1 - 0.4 = 0.6$. Since the auctioneer only uses 100 different words and all of the $100-20 = 80$ other words are equally likely, the probability of each of the other words is $0.6/80 = 0.0075$.

(a) In this example the discrete messages are the possible words. The entropy of the source in bits per message (word) is:

$$\sum_{i=0}^{100} -p_i \log_2(p_i)$$

$$= 20\times0.02 \log_2(0.02)+80\times0.0075 \log_2(0.0075)$$

$$\approx 6.5 \text{ bits/word}$$

(b) Since the auctioneer is speaking at a rate of 200 words per minute the information rate is $200/60 \times 6.5 \approx 21.6$ bps.

The information rate is also the minimum data rate required to transmit the information losslessly using the best possible compression. Thus the minimum data rate required, regardless of the type of compression is also 21.6 bps.

Question 2

To decompress the sequence `Call[4,2]tinH[5,2]` we would output the first sequence of letters (`Call`), then copy 2 letters starting 4 back from the current output (`Ca` resulting in having output `CallCa`), then copy the next four letters (`tinH` resulting in the output `CallCatinH`) then copy 2 letters starting 5 back from the current output (`at`) resulting in the final output: `CallCatinHat`.

Question 3

A particular communication system operates over a channel that is subject to fading (time-varying channel gain). The channel is "faded" about 10% of the time and during this time the bit error rate is 10^{-2}. The rest of the time the BER is 10^{-4}.

(a) Assume N bits are transmitted. The number of bits transmitted when the channel is faded is $0.1 \times N$ while $0.9 \times N$ bits are transmitted when the channel is not faded. The expected (average) number of errors will be $0.1 \times N \times 10^{-2} + 0.9 \times N \times 10^{-4} = 1.09N \times 10-3$. The expected fraction of bits in error (the BER) is thus $1.09 \times 10-3$.

(b) If the bit rate was R bps we would expected $R \times 24 \times 60 \times 60 \times 1.09 \times 10-3 \approx 94R$ bit errors.

(c) We need to find the probability that a frame will have an error (the Frame Error Rate or FER) given the BER.

There are many ($2^{128\times8} - 1$ in fact) ways that a frame could contain one or more errors but only one way that it could contain no errors. Rather than computing the probabilities of each of the possible frames that contain errors it's simpler to compute the probability that a frame is received without error. Then the probability that the frame is received with one or more errors is one minus that.

The probability of one bit being received correctly is $1 - P_e$ where P_e is the BER. For a frame to be received correctly the first bit needs to be correct AND the second bit needs to be correct AND ... If the errors and independent the probability that a frame with 1024 errors is received without any errors is the product of each of those probabilities or $(1 - P_e)^{1024}$.

The FER is thus $1 - (1 - P_e)^{1024}$.

When the channel is not faded $P_e = 10^{-4}$ and the FER is $1 - (1 - P_e)^{1024} \approx 0.1$

When the channel is faded $P_e = 10^{-2}$ and the FER is $1-(1-P_e)^{1024} \approx 1$

Question 4

It takes $(1024+64)\times 8\times 100\times 10^{-9}+2\times 20\times 10^{-6} \approx 910\ \mu$s to transmit a pair of frames (including the 20 μs gaps between frames).

In that time we receive $(1008+48)\times 8 = 8448$ bits of useful data (single-user, error-free channel).

Thus the throughput is $8448/910 \approx 9.3$ Mbps.

Question 5

The C program in Listing 1 contains a UTF-8 encoded string constant and produces the output shown in Listing 2.

Question 6

Listing 3 shows a C program that counts the number of ASCII characters and Unicode control points in the standard input. ASCII characters are single bytes with values between 0 and 127. Each UTF-8 sequence that is not ASCII begins with a byte whose MS two bits are 11 (subsequent bytes have the two MS bits set to 10). The program counts bytes that meet these two conditions and prints the totals when it sees the end of the file.

The result of running this program on the test input file is:

```
Found 13 ascii and 2 unicode sequences.
```

Question 7

The reflected pulse caused the voltage seen at the generator to drop so it must have been a negative voltage. At a short-circuit termination the sum of the incident and reflected waves must be zero so they must have opposite signs. Thus the diagram shows the effect of a reflection from a short circuit.

The two-way delay was 3 μs so the delay to reach the end of the transmission line must have been 1.5 μs. For an air dielectric the velocity of propagation is approximately the same as that in free space (3×10^8 m/s) so the length of the transmission line must be $ct = 3\times 10^8 \times 1.5\times 10^{-6} = 450$ m.

Question 8

The breakdown voltage of air is approximately 30 kV/cm and its dielectric constant is approximately 1.

Using the formula $P = V_{\text{rms}}^2/R$, a 1 MW signal on a 50 Ω transmission line will have an RMS voltage of $\sqrt{10^6\times 50} \approx 7.1$ kV.

For a sine wave the RMS voltage is $\frac{1}{\sqrt{2}}$ times the amplitude (peak voltage). Thus the peak voltage is 10 kV and the dielectric thickness (air gap) must be at least $\frac{10\text{ kV}}{30\text{ kV/cm}} = \frac{1}{3}$ cm[1].

We also know the equation for co-ax is:

$$Z_0 \approx \frac{60}{\sqrt{\varepsilon_r}} \ln\left(\frac{D}{d}\right)$$

where d and D are the center conductor and shield diameters (respectively) that we are trying to find.

The thickness of the dielectric is $(D-d)/2 = 0.33$ cm so we can substitute $D = 0.66+d$, $Z_0 = 50$ and $\varepsilon_r = 1$ in the above equation to get:

$$50 = 60 \ln\left(1 + \frac{0.66}{d}\right)$$

and solve for d:

$$d = \frac{0.66}{\exp\left(\frac{50}{60}\right) - 1} \approx 0.51 \text{ cm}$$

from which we can find $D = 0.66 + d \approx 1.2$ cm.

Question 9

For $d = 1\times 10^{-3}$, $W = 3\times 10^{-3}$ and $\varepsilon_r = 4$,

$$C = \frac{\varepsilon W l}{d} = \varepsilon 3l \quad \text{F}$$

and

$$L = \frac{\mu l d}{W} = \frac{\mu l}{3} \quad \text{H.}$$

Since for a transmission line:

[1] This assumes the electric field strength is constant everywhere between the inner coductor and shield. It is also too small a gap to use in practice.

```
#include <stdio.h>
#include <string.h>

main()
{
   char *p, *s = "   " ;
   printf ("%s takes %d bytes:", s, strlen(s)) ;
   for ( p = s ; *p ; p++ )
      printf (" %02x", (unsigned char) *p ) ;
   printf ("\n") ;

}
```

Listing 1: Sample Solution for Question 5.

```
埃德卡萨斯 takes 15 bytes: e5 9f 83 e5 be b7 e5 8d a1 e8 90 a8 e6 96 af
```

Listing 2: Output of Program in Listing 1.

```
#include <stdio.h>

main()
{
   int c, nascii=0, nunicode=0 ;
   while ( ( c = getchar() ) != EOF ) {
      if ( c <= 127 ) nascii++ ;
      if ( ( c & 0xc0 ) == 0xc0 ) nunicode++ ;
   }
   printf ("Found %d ascii and %d unicode sequences.\n", nascii, nunicode ) ;
}
```

Listing 3: Sample Solution for Question 6.

$$Z_0 = \sqrt{\frac{L}{C}}$$

we can substitute the values of L and C above to get:

$$Z_0 = \sqrt{\frac{\mu l}{3}\frac{1}{\varepsilon 3l}}$$

we are given $\varepsilon = \varepsilon_r \varepsilon_0 = 36 \times 10^{-12}$ F/m and $\mu = \mu_r \mu_0 = 1.3 \times 10^{-6}$ H/m so that $Z_0 = \frac{1}{3}\sqrt{\mu/\varepsilon} \approx 63\,\Omega$.