Assignment 1

Due Wednesday, October 28. Show your work. Submit your assignment using the appropriate dropbox on the course web site. Assignments submitted after the solutions are made available will be given a mark of zero.

Question 1

Consider an auctioneer speaking at a rate of 200 words per minute. Assume the auctioneer only uses 100 different words.

(a) What is the entropy of this source in bits per word if 20 of these words have a probability of 0.02, and the rest are equally likely?

(b) If you could compress the auctioneer's speech, what is the minimum data rate (in bps) required to transmit his/her speech without errors?

Question 2

The most common file compression technique is to replace a sequence of bytes that is repeated within a file with a reference to the previous occurrence.

In the following text, the notation [n,m] means to copy the *m* characters that appeared *n* bytes previously in the output into the output.

For example, `abcd[3,2]ef[4,3]` would be decompressed as: "`abcd bc ef bce`" where the spaces are inserted to mark the four output sequences and are not part of the output.

Decompress the following compressed sequence: `Call[4,2]tinH[5,2]`.

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) What is the long-term average bit error rate?

(b) How many bit errors would you expect per day?

(c) If the errors happen randomly, what is the probability that a frame with 128 bytes will have an error while the channel is not faded? What is the FER when the channel is faded?

Hints: Do not assume the average rate is the average of the rates. To find the FER, first find the probability that a bit is not in error and then the probability that all bits in a frame are not in error.

Question 4

A communication system transmits frames of two different lengths. One is 1024 bytes long, the second is 64 bytes long. Both frames are equally likely. These frame lengths include 16 bytes of overhead and so the frames contain 1008 and 48 bytes of useful data. A 20 μs gap must be left between each frame. If the system transmits frames one after the other and it takes 100 ns to transmit one bit, what is the throughput? Assume no errors and a single user.

Question 5

A C program can handle UTF-8 encoded strings just like ASCII strings since both are sequences of non-zero bytes terminated by a zero ("null") byte. However, the length of a UTF-8 encoded string is the number of bytes rather than the number of "characters" and language-specific operations such as case conversion may not work as expected.

The C program in Listing 1 contains a UTF-8 encoded string constant and will compile on most modern C compilers.

Use `http://translate.google.com` or the equivalent `bing.com` site to "translate" your name to Chinese, Japanese or Korean[1]. Change the string

[1] Note that this will probably not give you a name that is appropriate for other purposes. If you already have a name in a CJK language you may use that.

in the program to the resulting name in UTF-8 encoding. Compile and run the modified program and capture the displayed result.

Your answer should include the modified program listing and a screen capture of the output.

Hint: If the output of your program does not display properly on the console, redirect the output to a file as you did in lab1, read that file into a word processor or web browser and set the document's character encoding to UTF-8.

Question 6

Listing 2 shows a C program that counts the number of control characters in the standard input.

Modify the program to count both the number of ASCII characters and the number of Unicode code points in the standard input. Include a program listing and the result of running your program on the supplied test file (e.g. `tcc -run asg.c <asg1test.txt`).

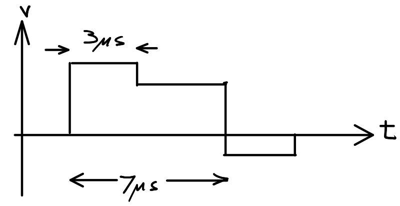
Hint: ASCII characters have the MS bit equal to zero. How can you identify the first byte of a UTF-8 sequence?

Question 7

In the lab you measured the length of a transmission line by observing the voltage at the source (generator) end of a transmission line when the remote end was terminated in a short circuit.

Length measurements can also be made by transmitting pulses. The transmitted pulse is reflected from the end of the transmission line and is seen arriving back at the generator after a delay. The voltage seen at the generator is the sum of the transmitted and reflected voltages.

The diagram below shows the voltage versus time at the source (generator) end of the transmission line made using air as the dielectric. What is the length of the transmission line? Is the termination a short-circuit or open-circuit?



Hint: If the end of the transmission line is shorted, what is the sum of the incident and reflected waveforms? If the end of the transmission line is open, what is the sum of the incident and reflected currents? In both cases, how much power is consumed by the termination?

Question 8

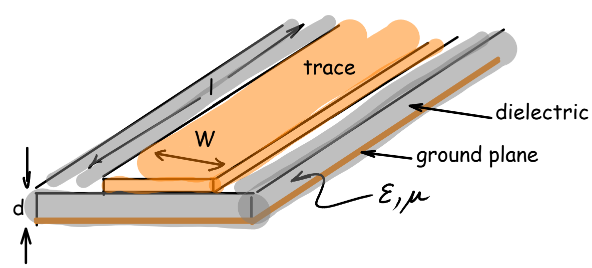
The breakdown voltage of air is approximately 30 kV/cm and its dielectric constant is approximately 1.

What are the diameters of the center conductor and shield of the thinnest air-dielectric 50 ohm coaxial cable capable of carrying a signal with a power of 1 MW without dielectric breakdown?

Hint: Start by finding the amplitude (peak voltage) of the signal ($P = V_{rms}^2/Z_0$) and use this to determine the minimum dielectric thickness.

Question 9

Microstrip is a common type of transmission line that is easily built on printed circuit boards (PCBs). It consists of a PCB trace separated from a ground plane layer by the PCB substrate:



There are equations that accurately predict the characteristic impedance of microstrip transmission lines. We can also make a (very) rough estimate by assuming the trace forms a parallel-plate capacitor and using an approximate equation for the inductance of a trace over a ground plane.

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

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

}
```

Listing 1: Sample C Program for Question 5.

```c
#include <stdio.h>

main()
{
   int c, n = 0 ;
   while ( ( c = getchar() ) != EOF ) {
     if ( c < 32 ) n++ ;
   }
   printf ("Found %d control characters\n", n ) ;
}
```

Listing 2: Sample C Program for Question 6.

The capacitance of a parallel-plate capacitor is given by:

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

where W is the plate width, l is the plate length, d is the dielectric thickness and ε is the absolute (not relative) permittivity ($\varepsilon = \varepsilon_r \varepsilon_0$). $\varepsilon_0 \approx 9$ pF/m.

The inductance of a PCB trace is approximately:

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

where l and W are the trace length and width respectively, d is the substrate thickness and μ is the absolute (not relative) permeability. Similar to the dielectric constant, $\mu = \mu_r \mu_0$. For a PCB substrate $\mu_r = 1$. $\mu_0 \approx 1.3\ \mu$H/m.

Estimate the characteristic impedance of a microstrip transmission line with a trace width of 3 mm over a substrate of 1 mm thickness and $\varepsilon_r = 4$.