ELEX 3525 : Data Communications
2018 Fall Term

Assignment 1

Due Tuesday, October 9, 2018. Submit your assignment using the appropriate Assignment folder on the course web site. Assignments submitted after the solutions are made available will be given a mark of zero. **Show how you obtained your answers.**

Question 1

The CSV file, `asg1_1.csv` on the course web site contains the words and letters in a news story along with the number of times each word and each letter occurs in the story. Case, non-alphabetic characters and punctuation have been ignored.

How many bits of information are contained in the story if you treat:

(a) each word as a message,

(b) each character as a message, and

(c) each character as having 8 bits of information?

Hint: use a spreadsheet.

Question 2

Compute the average throughput per user for a communication system that transmits frames at a rate of 2 Mb/s with the following overhead per frame:

(a) a preamble of 10 bytes and a header of 7 bytes per frame,

(b) 12 bytes of parity for every 64 bytes of payload (rounded up),

(c) a guard time of 8 μs following each frame.

Half of the frames are 64 bytes long and half are 1500 bytes long. The channel is shared by 10 users and there are no errors or retransmissions. Give your answer in kb/s.

Question 3

The following sequence of bytes (values in hex) are UTF-8 encoded characters mixed in with some bytes that are not part of valid UTF-8 encodings.

```
E1 A2 84 BE E3 81 AE 45 8A D0 B7
```

(a) Identify the bytes that are and are not parts of valid UTF-8 encodings by ignoring invalid bytes. Assume the first byte is correct.

(b) What are the Unicode code points corresponding to each valid UTF-8 encoded character? Give your answers in hex.

(c) What are the names of the corresponding characters? (*Hint: see unicode.org code charts for the character tables*).

Question 4

Data is transmitted over a serial interface using 8 data bits and one parity bit for error detection. The probability of a data or parity bit being received in error is 10^{-6} and is the same for each bit. You can ignore errors on start and stop bits.

(a) What is the probability that the first data bit is in error?

(b) What is the probability of exactly one bit in a character being in error?

(c) What is the probability of exactly two bit errors in one character?