

Problem Set #7

ENEE 426, Spring 2008

Discussed in Class, Tuesday, May 6

Complete the following problems:

1. Compute the entropy of the following random variables:
 - a. The roll of a die – $\Pr[X=\{1, 2, \dots, 6\}] = 1/6$
 - b. A biased coin – $\Pr[X=\text{heads}]=3/4$; $\Pr[X=\text{tails}]=1/4$
2. Use an English corpus, such as a large text file containing a dictionary or novel (these can be readily found on the Internet) to build a probability table for English characters. The easiest way to do this is count the number of times characters occur in the corpus and then divide by the total number of characters. Consider a normalized alphabet containing lower-case letters and a space. Use this probability table to compute the entropy of a random variable over the English alphabet. Compare this to the maximal entropy of a 27-character alphabet ($\log_2(27)=4.7$).
3. Develop a Huffman Code to optimally encode the normalized alphabet from #2.
4. Suggest a way to achieve better compression of English text over using Huffman codes on a per-character alphabet?
5. Design an XML Schema to represent symbol mappings and data for a Huffman-encoded message. How long does the message have to be to overcome the overhead of transmitting the XML Schema?
6. What properties of WiFi make MAC-layer encryption more necessary than in a wired Ethernet network?
7. Describe possible attacks that could be occurring if you visit a website whose certificate is invalid?