15-211: Assignment 4	Theory Questions
Due: June 13, 2008 in class	

These questions count for 15% of the homework grade. Please hand in your answers in lecture on Friday.

1 Huffman vs. LZW (8 pts)

(a) (2 points) Quantra Disk Company has come up with a hard drive that can store 4 possible values in each cell (0,1,2,3) rather than the typical (0,1). How can you modify the Huffman Encoding algorithm to take advantage of this new technology?

(b) (2 points) You initialize your dictionary for LZW compression with 256 elements. Then, you read in 69 bytes from a file you want to compress. What is the minimum and maximum number of entries in the dictionary?

Minimum: 256 + _____

Maximum: 256 + _____

(c) (4 points) Suppose you're compressing a file consisting of a's,b's,c's, and d's with LZW. Your initial dictionary thus has only four elements (a=0, b=1, c=2, d=3). You want to compress the following string:

$$\text{baadcaadaad}$$

Show the end dictionary after you run LZW:

a=0	
b=1	
c=2	
d=3	

What does it compress to?

2 Burrows-Wheeler (7 pts)

(a) (1 point) Encode the following string of numbers using Huffman encoding:

01212312

What is the length of the result in bits?

(b) (2 points) Perform the Burrows-Wheeler transform on the original string:

01212312

What is the resulting string?

(c) (2 points) Perform the Move-To-Front transform on your string from the previous part. The initial dictionary is as follows:

[0,1,2,3]

What is the resulting string?

(d) (2 points) Why does the Burrows-Wheeler transform followed by Move-To-Front generally result in smaller Huffman compressed files?