Due: Wednesday, October 7

This assignment is worth 15% of final grade

1. Write down the initial of your first name and the first three letters of your last name. For example, for Allan Borodin, I would be writing abor. Encode this shortened name (e.g. abor) by encoding a as 01, b as 02, ..., z as 26. If your last name has less than three letters, then use a blank symbol encoded by 00. That is, abor would be encoded as 01021518 and Allan Bo would be shortened to abo encoded as 01021500.

2. Let ID denote the encoded version of your shortened name. That is, ID an 8 digit decimal number (7 digits if we don't count a leading 0). (Given the class list I believe this is a unique identifier for every student in the class.) How many possible values can there be for an ID assuming that every 4 symbol name is possible? Note: There are no three symbol IDs.

3. Now we want to introduce an example of an important idea (maybe a great idea) called hashing. Let $h(ID) = ID(mod\ 83)$. By "mod 83", I mean divide by 83 and take the remainder. For example, using my ID, we would get $1021518(mod\ 83) = 37$. That is, my unique ID (in this class) is being hashed or mapped onto a much smaller number (i.e. a number between 0 and 82). Note: We are discussiing why hashing is a useful idea in our seminar. We will also go over a few basic probability concepts in class.

4. Suppose we assume that this is a perfect hash function (which it is not) in the following probabilistic "balls and bins" sense: We think of randomly (with uniform probability $\frac{1}{83}$) and independently throwing each ID (as if it were a ball) at one of 83 possibles bins. Suppose there were only 4 students in the classs. Calculate the probability that (at least) two students will be hashed to the same hash value. Hint: First calculate the probability that all 4 students get a unique hash value.

5. We have approximately 30 students in CSC196 (each having a unique ID) and 83 possible hash values. Once again, assume the hash function is a perfect hash function. How likely do you think it is that there will be two students with the same hash value? I am not asking for a detailed probabilistic analysis, just your intuitive explanation of how likely this is. Hint: You may want to look up the Birthday Paradox.