**Reading for Lecture 12:**

**Read:** *D is for Digital* Part II Software: Intro and Chapter 4 Algorithms (pp. 51-63)

**Lecture 11: Exercises**

1. Recalling that there are about 3.2 billion (3,200,000,000 = 3.2 x 10\*\*9)) base pairs in the human genome and each base pair represents one of four possibilities, calculate the storage required to record a complete human genome, in bits. Show your work.

2. A typical CD holds about 650 megabytes (= 650,000,000 bytes = 6.5 x 10 \*\*8 bytes). How many CDs would you need to store the information in your genome?

3. At the top of p. 174 in *D is for Digital*, Kernighan displays a pair of fraudulent emails that the author received, ostensibly from a friend in need, but actually from a scammer. The article by Stajano and Wilson names and explains seven “principles” used by scammers. For each of these principles, explain whether you think it is present or absent in this particular scam, and explain your answer.

4.A. Suppose you have to search an unsorted list of 1000 names for a particular name. If you just start at the beginning and begin checking one name after another, comparing each name with the target name, how many comparisons will you have to make, on the average, before you find the match?

4.B. Now suppose the list of 1000 names is sorted alphabetically and instead of just starting at the beginning, you use a binary search to locate the match. How many comparisons will you have to make, at most, to find the matching name?

5. Hashing. You can find a calculator for the Secure Hash Algorithm – 256 (SHA-256) here: <http://www.xorbin.com/tools/sha256-hash-calculator>

Use the calculator to generate a SHA-256 hash for the Gettysburg address, using the text below (cut and paste the text). The resulting has will have 64 hexadecimal digits.

Four score and seven years ago our fathers brought forth on this continent, a new nation, conceived in Liberty, and dedicated to the proposition that all men are created equal. Now we are engaged in a great civil war, testing whether that nation, or any nation so conceived and so dedicated, can long endure. We are met on a great battle-field of that war. We have come to dedicate a portion of that field, as a final resting place for those who here gave their lives that that nation might live. It is altogether fitting and proper that we should do this. But, in a larger sense, we can not dedicate -- we can not consecrate -- we can not hallow -- this ground. The brave men, living and dead, who struggled here, have consecrated it, far above our poor power to add or detract. The world will little note, nor long remember what we say here, but it can never forget what they did here. It is for us the living, rather, to be dedicated here to the unfinished work which they who fought here have thus far so nobly advanced. It is rather for us to be here dedicated to the great task remaining before us -- that from these honored dead we take increased devotion to that cause for which they gave the last full measure of devotion -- that we here highly resolve that these dead shall not have died in vain -- that this nation, under God, shall have a new birth of freedom -- and that government of the people, by the people, for the people, shall not perish from the earth.

SHA-256 hash of the above (64 hexadecimal digits):

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Now change the initial character ‘F’ to a lower case ‘f’ and recompute the hash.

Write the second hash under the first one.

How many hex digits match up in the two hashes?