CSCI 340
Data Structures and Algorithms
Spring 2013

Project 0 – Needle in a Haystack

Due date: February 13th Prior to Class

Background

The first assignment has multiple purposes. Brush up on your basic Java I/O, show me what you know, practice doing a bit of bit manipulation, check that you can get to the class directories for test files, test the lab computer's submit function, and see if you can follow directions (and how well you can follow them.)

Your Assignment

Part1

Find a needle in a haystack, actually more like find the word needle in a file that I supply, called the haystack. The word needle may appear in any cAsE, but it will always be alone on a line. Your task is to search through the given file or files and tell me which file and line has the needle, you can stop after the first needle. You should output filename : linenumber

Part 2

Now I want you to find a binary representation of a needle. The null byte (eight 0s) will not often be found in a plain text file, but I want you to find the first occurrence of 8 sequential bits all set to zero. For this part, you only have to search one file and you can stop at the first needle. You should output the two characters whose bits are combined to make eight sequential zeros (e.g. lowercase p followed by a tab character). Think about this for a minute why would it always be only two letters that create 8 sequential 0s and not 3 letters?

Note: one way to do this would be to find all the combinations of ascii characters that would have 8 zeros when placed next to each other – DON'T do it this way.

Two big hints here, one you are probably going to want to use the java.util.BitSet which is a class for holding collections of bits. One feature it doesn't have is how to take an array of java bytes and convert it to a BitSet (Which this function does – although be careful if you send it more than 1 byte at a time which bit of which byte ends up at the top index? The bottom?)

```
public static BitSet fromByteArray(byte[] bytes) {
        BitSet bits = new BitSet();
        for (int i = 0; i < bytes.length * 8; i++) {
                if ((bytes[bytes.length - i / 8 - 1] & (1 << (i % 8))) > 0) {
                        bits.set(i);
                }
        }
        return bits;
```

}

Sample haystacks are available at: `/home/student/Classes/Cs340/Assignment0`

Good software engineering is expected. Use lots of comments, appropriate indentation, etc. when writing your code. Also good function names – input parameter names, etc.

What to Submit – READ CAREFULLY

Part 1

A single .java file named Needle.java that takes arguments from the command line and prints back to the console.

- Use package name **needle** and class **Needle** for your main class
- I will run the following commands on your submission, for full credit this should generate output:
  - javac –d . Needle.java (that's a dash d with a period after it)
  - java needle/Needle filename1 filename2 filename3 etc….

Part 2

A single .java file named NeedleB.java that takes arguments from the command line and prints back to the console.

- Use package name **needle** and class **NeedleB** for your main class
- I will run the following commands on your submission, for full credit this should generate output:
  - javac –d . NeedleB.java (that's a dash d with a period after it)
  - java needle/NeedleB filename

How to submit

- Use the submit tool (submit 340 filename) to submit each java file (2 files total) this works only from a lab machine or lab ssh session.
- You don't have to work on the lab machine, but I will only accept work submitted through the submit tool – **do not** email me your .java file. If you have problems please let me know.