Command-line Bookkeeping (v1)

Purpose

At the end of this laboratory assignment, you should be able to use the command line tools `javac` and `java` to compile and run a JAVA program.

You should also be able to use basic UNIX command ideas, like pipes.

Due Date

The completed lab assignment is due Friday, 2018-01-19 *by the beginning of lecture.*

Information about Command Line Programming

This section contains background information that you will use later in the lab assignment.

The command line and `public static void main(...`

Your laboratory instructor should show you how to compile java programs with the "`javac`" command, and how to run them with "`java`".

If you write a `HelloWorld.java` file

```
public class HelloWorld {
    public static void main(String [] args) ...
```

and run

```
% java HelloWorld Fred Gertrude
```

the strings `"Fred"` and `"Gertrude"` will be stored in `args[0]` and `args[1]`. *Command line information is communicated from the command-line to your program via the argument of the* `main` *method.*

`System.in`, `System.out` and the command line

UNIX command-lines let you redirect where your programs read from and where they write to. If you run

---
```
% java HelloWorld > myOut.txt
```
---

information written by `HelloWorld` gets stored in the file `myOut.txt`. Note that `System.out` doesn't always connect with user-visible output!

Similarly, if you run

```
% java HelloWorld < myIn.txt
```

`HelloWorld` will read information from the file `myIn.txt` rather than from the console. Note that `System.in` doesn't always connect with the user's keyboard!

You can also do this by running

```
% cat myIn.txt | java HelloWorld
```

In this case the "`cat`" command reads the file "`myIn.txt`" and copies it to its `System.out`, but the '`|`' causes the output of `cat` to become the input of `java HelloWorld`. Using multiple pipe symbols, you can link three or more programs into a *pipeline*.

"Bookkeeper's" friends

Not many words in English contain three or more consecutive pairs of letters; "book-keeper" is one of the few examples. However, there are a fair number that contain two pairs: "coffee" being an example.

In Finnish, there are more words with lots of pairs. For instance, "hyppyyttää" (which may mean "dance" according to Google translate), and "periaatteellinen" (principled).

Your overall task for this laboratory assignment is to write a JAVA program whose input is an integer n and the name of a file. The program should then read the file, which typically has one word per line, and print to standard out ("`System.out`") all of the words that have n or more consecutive pairs.

The program to write

⇒ Write a method (possibly `static`) whose arguments are

- an integer specifying the minimum number of consecutive pairs to accept;
- a `java.util.Scanner` that provides words via `.nextLine()`; and
- a `java.io.PrintWriter` object where words with the correct number of pairs are written.

When called, the routine should read each possible words from the scanner, check to see if it has an appropriate number of consecutive pairs, and, if it does, write the word to the `PrintWriter`.

This may (probably should) have sub-methods; it is the signature which is important.

⇒ Write a **public static void main(String [] args)** that utilizes JAVA's command line ability to accept command-line arguments.

- The first command-line argument, if present, should be the number of consecutive pairs to search for. If no number is supplied, the default should be 3.
- The second command-line argument, if present, should be the name of a file to read from. If no second argument is supplied, the program should read lines from `System.in`.
- The third command-line argument, if present, should be the name of a file to write to. If no third argument is supplied, the program should write to `System.out`.

⇒ Use the method that you wrote in the first part together with your main method to implement a program that can search a file (possibly standard input) for a given number of pairs of words. The words that match the criterion should be saved to a file, or by default, displayed on `System.out`.

Think very hard about the required logic, and code as cleanly as you can.

Command-line programming tasks

Regardless of where or how you develop your code, transfer it to a lab machine in 8-457 and run it there.

⇒ *Ask your laboratory instructor how to create script files. For this assignment, the file that you upload to* Blackboard *should be a cleaned script file.*

⇒ Use a text editor of some kind to create a very small list of words (possibly imaginary words) that you can use to test your program. Test your program by using your small list of words.

⇒ Download a plain-text file of Finnish words. (I recommend `https://raw.github-usercontent.com/hugovk/everyfinnishword/master/kaikkisanat.txt`, which you can find on `GitHub` (search for "every Finnish word").)

⇒ Using a command-line similar to

```
java Bookkeeper 4 kaikkisanat.txt
```

print out a list of Finnish words that have at least four consecutive pairs of letters.

⇒ Learn what the UNIX `wc` command does. Use your `Bookkeeper` program, a command-line pipe and `wc` to count the number of Finnish words that have at least three consecutive pairs of letters.

⇒ Using a command-line similar to

```
cat /usr/share/dict/words | java Bookkeeper 3
```

print out a list of English words that have at least three consecutive pairs of letters.

⇒ Determine the number of English words that have two or more pairs of consecutive pairs of letter.