

Programming Languages and Techniques (CIS120)

Lecture 30

April 9, 2014

Histogram Demo

Announcements

- HW09: Spellchecking is available now
 - Due Tuesday, April 15th at 11:59:59pm
 - Practice with Collections and IO libraries
- Today and Friday's lecture: learn about IO & collections libraries *through example* via in-class design exercise
 - Read Chapter 29
 - Ask questions!
 - No clickers on Friday
- Will return to exceptions and talk more about IO *after* we finish the design exercise
- Exam solutions posted, view exams in Levine 308

Poll

How many of these these classes have you used before CIS 120 (all part of the Java standard library)?

- Scanner
- Reader
- InputStream (e.g. System.in)
- FileReader
- BufferedReader
- Something else from java.io?

Design Example: Histogram.java

A design exercise using java.io and
the generic collection libraries

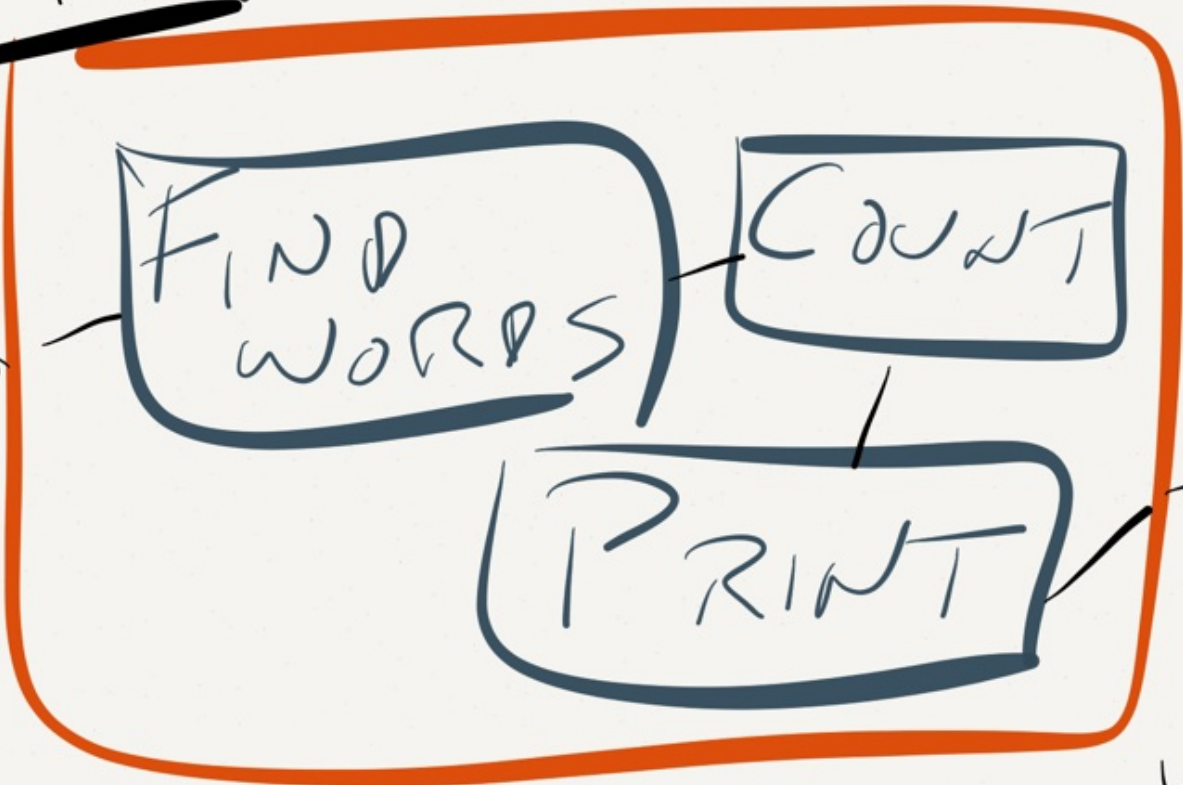
Problem Statement

Write a command-line program that, given a filename for a text file as input, calculates the frequencies (i.e. number of occurrences) of each distinct word of the file. The program should then print the frequency distribution to the console as a sequence of “word: freq” pairs (one per line).

Histogram result:

The : 1	each : 1	line : 2	should : 1
Write : 1	file : 2	number : 1	text : 1
a : 4	filename : 1	occurrences : 1	that : 1
as : 2	for : 1	of : 4	the : 4
calculates : 1	freq : 1	one : 1	then : 1
command : 1	frequencies : 1	pairs : 1	to : 1
console : 1	frequency : 1	per : 1	word : 2
distinct : 1	given : 1	print : 1	
distribution : 1	i : 1	program : 2	
e : 1	input : 1	sequence : 1	

TEXT FILE



PRINTED HISTOGRAM

Decompose the problem

- Sub-problems:
 1. How do we iterate through the text file, identifying all of the words?
 2. Once we can produce a stream of words, how do we calculate their frequency?
 3. Once we have calculated the frequencies, how do we print out the result?
- What is the interface between these components?
- Can we test them individually?

Interactive Demo

Histogram.java and WordScanner.java