

CS 146

Data Structures and Algorithms

Summer Semester 2015

Department of Computer Science
San José State University
Instructor: Ron Mak

Assignment #1

Assigned: Tuesday, June 2
Due: Monday, June 8 at 11:59 pm
100 points max

Text search

The purpose of this assignment is to practice determining the scalability of an algorithm.

Download the complete text of *War and Peace* as an ASCII file from <http://www.cs.sjsu.edu/~mak/CS146/assignments/1/WarAndPeace.txt>
The file contains over 65,000 lines and over a half million words.

Write a Java program to search for the following names in the text:

- Makar Alexeevich
- Joseph Bazdeev
- Boris Drubetskoy

For each occurrence of each name, print:

- The starting line number (first line is 1).
- The starting character position (first position is 1).
- The name.

Example output:

LINE	POSITION	NAME
19949	1	Boris Drubetskoy
21953	2	Makar Alexeevich
22173	9	Boris Drubetskoy
46612	19	Joseph Bazdeev

A name can be split across two consecutive lines. More than one name can be on a line. You must print the names in the order that they appear in the text.

Print how long your program runs:

```
long start = System.currentTimeMillis();
/* do everything here */
long elapsed = System.currentTimeMillis() - start;
```

Run your program several times and pick the median time.

Teamwork

You may work individually as a team of one, or you can partner with another student as a team of two.

You can be on only one team at a time. If you partner with someone, both of you will receive the same score for this assignment. You'll be able to choose a different partner or work alone for subsequent assignments.

What to turn in

Create a zip file containing:

- Your Java source files.
- A sample output file. Use output redirection, or cut-and-paste into a text file.
- A short report (at most 3 pages) describing your algorithm and how well it will scale with respect to the length of the text and the number and lengths of the names to search.

Name the zip file after yourself or yourselves.

Examples: `smith.zip`, `smith-jones.zip`

Each team should email the zip file to ron.mak@sjsu.edu. Your subject line must be:

```
CS 146 Assignment #1 Your name(s)
```

Example:

```
CS 146 Assignment #1 Mary Smith & John Jones
```

If you work with a partner, you should email only one assignment between the two of you. Whoever emails the assignment should CC the partner so that when I send you your team score, I can just do a "Reply all".