Java Assignment 3

This assignment is about parallelizing a program in two different ways.

You will need the following files (downloadable from iLearn2):
1) WebPage.java
2) Program.java

The class WebPage represents a web page, and have
a) a constructor,
b) a method download, which performs a fake download of text from a webpage (it just generates a string of nonsense words),
c) a method analyze, which counts occurrences of all different words occurring in the downloaded text,
d) a method categorize, which performs a categorization of the text based on the word with most occurrences in the text,
e) a method toString, which returns a string representation of the categorized webpage.

You are **not** allowed to modify the class WebPage.

The class Program performs the task of downloading a number of webpages, analyzing, categorizing and presenting these webpages in a sequential way. You **should** create variants of the class Program as described below. If you like to create some new classes, then do that as inner classes within the program class.

Basic task (grade C/D/E/F)

Create two programs Program1 and Program2 as specified below, and answer the two questions Question1 and Question2 below, in a textfile Answers.txt?

Program1 (in package paradis.assignment3) should parallelize the execution of Program by using the **producer-consumer pattern** and the **common ForkJoinPool** (thread pool). The different stages in the working process (initializing, downloading, analyzing, categorizing, presenting result) should be **separated** using the producer-consumer pattern. Let the number of threads of the thread pool be the automatically set value. The program (Program1) should be threadsafe and as performant as possible.

Program2 (in package paradis.assignment3) should parallelize the execution of Program by using **parallel Stream**, which implicitly uses the **common ForkJoinPool** (thread pool). The different stages in the working process (initializing, downloading, analyzing, categorizing, presenting result) should be **connected** using the **Stream** interface. Let the number of threads of the thread pool be the automatically set value. The program (Program2) should be threadsafe and as performant as possible.

Question1: Which of the two programs, Program1 and Program2, has best performance, or are the performance of the two programs about the same?

Question2: Is there any difference in which order different work tasks are performed in the two different solutions, Program1 and Program2?

Advanced task (grade A/B/C/F)

Solve the basic task as described above, implement a class MyExecutor and create a program Program3 as specified below, and answer the question Question3 below in the textfile Answers.txt.

MyExecutor (in package paradis.assignment3) should implement the interface ExecutorService, and it should manage a pool of threads where the programmer should be able to set the number of threads. In the implementation of MyExecutor you are not allowed to use any predefined thread-pool-class in Java. You only need to implement the methods of ExecutorService that you use, the rest of the methods could just throw UnsupportedOperationException.

Program3 (in package paradis.assignment3) should parallelize the execution of Program by using the **producer-consumer pattern** in the same way as Program1, but instead of the common ForkJoinPool it should use your class **MyExecutor**. Set the number of threads so you get the best performance. Otherwise, the program (Program3) should be threadsafe and as performant as possible.

Question3: Which of the two programs, Program1 and Program3, has best performance, or are the performance of the two programs about the same?

Submission

Your submission should be a **zip-file** called **Assignment3.zip** containing all the files as described above, and it should be submitted to the appropriate place in NEXTiLearn.

Good luck!