

XINHUI ZHAO (001560851)

Program Structures & Algorithms

Fall 2021

Assignment No. 5

◉ Task (List down the tasks performed in the Assignment)

Your task is to implement a parallel sorting algorithm such that each partition of the array is sorted in parallel. You will consider two different schemes for deciding whether to sort in parallel.

1. A cutoff (defaults to, say, 1000) which you will update according to the first argument in the command line when running. It's your job to experiment and come up with a good value for this cutoff. If there are fewer elements to sort than the cutoff, then you should use the system sort instead.
2. Recursion depth or the number of available threads. Using this determination, you might decide on an ideal number (t) of separate threads (stick to powers of 2) and arrange for that number of partitions to be parallelized (by preventing recursion after the depth of $\lg t$ is reached).
3. An appropriate combination of these.

You must prepare a report that shows the results of your experiments and draws a conclusion (or more) about the efficacy of this method of parallelizing sort. Your experiments should involve sorting arrays of sufficient size for the parallel sort to make a difference. You should run with many different array sizes (they must be sufficiently large to make parallel sorting worthwhile, obviously) and different cutoff schemes.

◉ Relationship Conclusion: (For ex : $z = a * b$)

Cutoff: 100,000

Threads number: 4

◉ Evidence to support the conclusion:

1. **Output (Snapshot of Code output in the terminal)**

```
ass5 src main java edu neu coe info6205 sort par Main main
Project
  > idea
  > logs
  > src
    > main
      > edu.neu.coe.info6205
        > balsearchtree
Run: Main
  "/Applications/IntelliJ IDEA CE.app/Contents/jbr/Contents/Home/bin/java" ...
Degree of parallelism: 7
cutoff: 50000 20times Time:2632ms
cutoff: 100000 20times Time:1774ms
cutoff: 150000 20times Time:1836ms
cutoff: 200000 20times Time:1842ms
cutoff: 250000 20times Time:1858ms
cutoff: 300000 20times Time:1895ms
cutoff: 350000 20times Time:1877ms
cutoff: 400000 20times Time:1982ms
cutoff: 450000 20times Time:1897ms
cutoff: 500000 20times Time:1871ms
cutoff: 550000 20times Time:1809ms
cutoff: 600000 20times Time:1959ms
cutoff: 650000 20times Time:1849ms
cutoff: 700000 20times Time:1839ms
cutoff: 750000 20times Time:1848ms
cutoff: 800000 20times Time:1854ms
cutoff: 850000 20times Time:1846ms
cutoff: 900000 20times Time:1861ms
cutoff: 950000 20times Time:1853ms
Process finished with exit code 0
```

When the array.length is 2,000,000, using the default ForkJoinPool and then running the cutoff from 50,000 to 950,000, it shows that the best case is around 100,000. So I changed the times to 100 and running the cutoff from 10,000 to 150,000, and actually, it shows that 100,000 is the best cutoff, as below.

```
ass5 src main java edu neu coe info6205 sort par Main main
Project
  > idea
  > logs
  > src
    > main
      > edu.neu.coe.info6205
        > balsearchtree
        > bqs
        > codelength
        > coupling
        > equable
        > functions
        > graphs
        > greedy
        > hashtable
Run: Main
  "/Applications/IntelliJ IDEA CE.app/Contents/jbr/Contents/Home/bin/java" ...
Degree of parallelism: 7
cutoff: 10000 100times Time:18294ms
cutoff: 20000 100times Time:9597ms
cutoff: 30000 100times Time:9444ms
cutoff: 40000 100times Time:9128ms
cutoff: 50000 100times Time:8773ms
cutoff: 60000 100times Time:9265ms
cutoff: 70000 100times Time:8525ms
cutoff: 80000 100times Time:8493ms
cutoff: 90000 100times Time:8424ms
cutoff: 100000 100times Time:8420ms
cutoff: 110000 100times Time:8465ms
cutoff: 120000 100times Time:8436ms
cutoff: 130000 100times Time:8466ms
cutoff: 140000 100times Time:8706ms
cutoff: 150000 100times Time:8984ms
Process finished with exit code 0
```

And I changed the array.length to 1,000,000, 1,500,000, 2,500,000 and 3,000,000, it also seems that 100,000 is the best value of cutoff.

The top screenshot shows the following code in `Main.java`:

```
15  * This code has been fleshed out by @yuguo @yao. Thanks very much.
16  * 7000 tidy it up a bit.
17  */
18  public class Main {
19
20      public static void main(String[] args) {
21          processArgs(args);
22          System.out.println("Degree of parallelism: " + ForkJoinPool.getCommonPoolParallelism());
23          Random random = new Random();
24          int[] array = new int[1000000];
25          ArrayList<Long> timeList = new ArrayList<>();
26          for (int j = 1; j < 16; j++) {
27              ParSort.cutoff = 10000 * j;
28              //ParSort.cutoff = 100000;
29              //for (int j = 1; j < 20; j++){
```

The output shows the following results:

cutoff	100times Time
100000	5180ms
200000	4594ms
300000	4656ms
400000	4454ms
500000	4355ms
600000	4572ms
700000	4531ms
800000	4382ms
900000	4461ms
1000000	4420ms
1100000	4470ms
1200000	4470ms
1300000	4892ms
1400000	5096ms
1500000	4848ms

The bottom screenshot shows the following code in `Main.java`:

```
18  public class Main {
19
20      public static void main(String[] args) {
21          processArgs(args);
22          System.out.println("Degree of parallelism: " + ForkJoinPool.getCommonPoolParallelism());
23          Random random = new Random();
24          int[] array = new int[1500000];
25          ArrayList<Long> timeList = new ArrayList<>();
26
27          for (int j = 1; j < 16; j++) {
28              ParSort.cutoff = 10000 * j;
29              //ParSort.cutoff = 100000;
30              //for (int j = 1; j < 20; j++){
```

The output shows the following results:

cutoff	100times Time
100000	7410ms
200000	6517ms
300000	7015ms
400000	7176ms
500000	6922ms
600000	6931ms
700000	6926ms
800000	6937ms
900000	6932ms
1000000	6779ms
1100000	6799ms
1200000	6838ms
1300000	6800ms
1400000	7021ms
1500000	7186ms

Project: edu.neu.coe.info6205.sort.par

```
public static void main(String[] args) {
    processArgs(args);
    System.out.println("Degree of parallelism: " + ForkJoinPool.getCommonPoolParallelism());
    Random random = new Random();
    int[] array = new int[250000];
    ArrayList<Long> timeList = new ArrayList<>();

    for (int j = 1; j < 16; j++) {
        ParSort.cutoff = 10000 * j;
        //ParSort.cutoff = 100000;
        //for (int j = 1; j < 20; j++){
            //...
        }
    }
}
```

Run: Main

"/Applications/IntelliJ IDEA CE.app/Contents/jbr/Contents/Home/bin/java" ...

Degree of parallelism: 7

cutoff	100times Time
10000	13772ms
20000	12090ms
30000	12402ms
40000	11624ms
50000	11328ms
60000	11313ms
70000	11287ms
80000	11839ms
90000	11024ms
100000	11009ms
110000	11026ms
120000	11034ms
130000	11022ms
140000	11396ms
150000	11215ms

Process finished with exit code 0

Project: edu.neu.coe.info6205.sort.par

```
* This code has been fleshed out by Yiyao Qiao. Thanks very much.
* 7000 tidy it up a bit.
*/
public class Main {

    public static void main(String[] args) {
        processArgs(args);
        System.out.println("Degree of parallelism: " + ForkJoinPool.getCommonPoolParallelism());
        Random random = new Random();
        int[] array = new int[500000];
        ArrayList<Long> timeList = new ArrayList<>();

        for (int j = 1; j < 16; j++) {
            ParSort.cutoff = 10000 * j;
            //ParSort.cutoff = 100000;
        }
    }
}
```

Run: Main

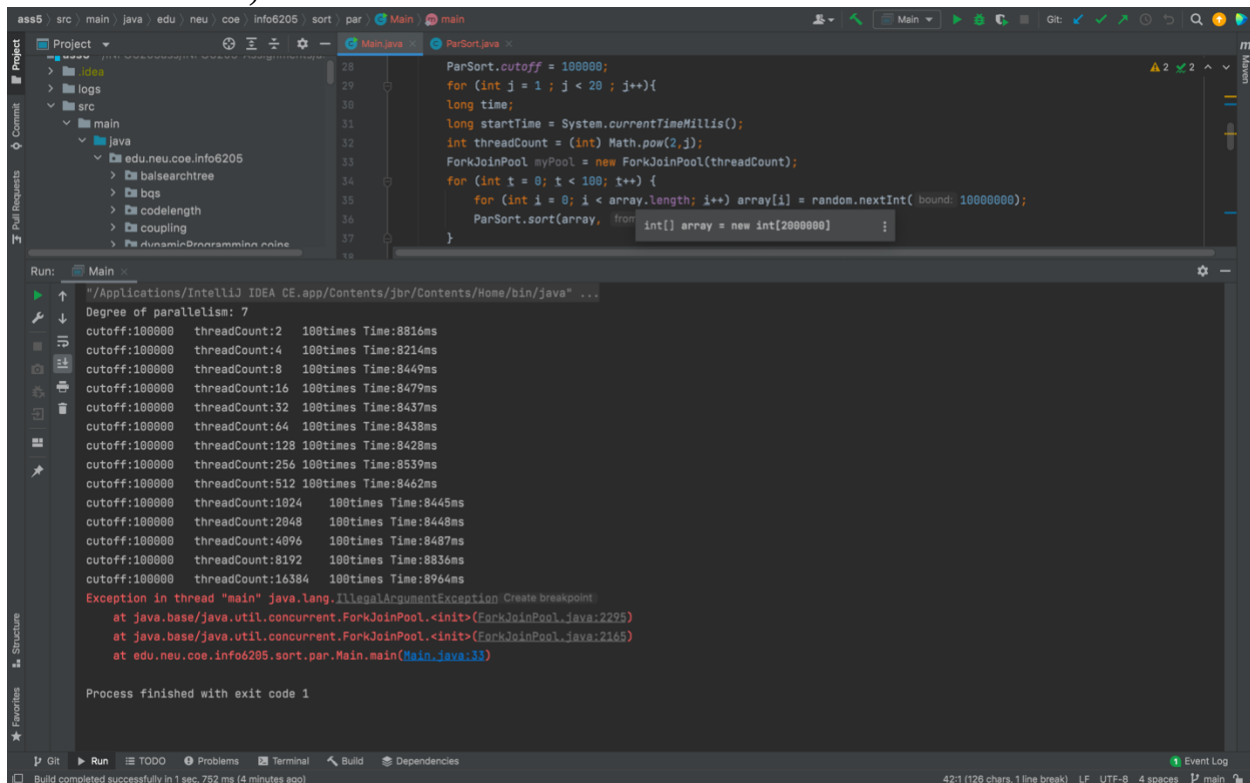
"/Applications/IntelliJ IDEA CE.app/Contents/jbr/Contents/Home/bin/java" ...

Degree of parallelism: 7

cutoff	100times Time
10000	17099ms
20000	15637ms
30000	14699ms
40000	14786ms
50000	14119ms
60000	13999ms
70000	14137ms
80000	14085ms
90000	14058ms
100000	13514ms
110000	13516ms
120000	13509ms
130000	13483ms
140000	13568ms
150000	13487ms

Process finished with exit code 0

When the cutoff is 100,000 and array.length is 2,000,000, running the thread count from 2^1 to 2^{20} , and the max value of thread count is 2^{14} , We can find the best value of thread count is 4.



The screenshot shows the IntelliJ IDEA IDE. The top editor displays a Java file named `ParSort.java` with the following code:

```
28 ParSort.cutoff = 100000;
29 for (int j = 1; j < 20; j++){
30     long time;
31     long startTime = System.currentTimeMillis();
32     int threadCount = (int) Math.pow(2, j);
33     ForkJoinPool myPool = new ForkJoinPool(threadCount);
34     for (int t = 0; t < 100; t++) {
35         for (int i = 0; i < array.length; i++) array[i] = random.nextInt( bound: 10000000);
36         ParSort.sort(array, from: int[] array = new int[2000000]
37     }
38 }
```

The bottom editor shows the output of the program's execution. It starts with the command `"/Applications/IntelliJ IDEA CE.app/Contents/jbr/Contents/Home/bin/java" ...` and the output:

```
Degree of parallelism: 7
cutoff:100000 threadCount:2 100times Time:8816ms
cutoff:100000 threadCount:4 100times Time:8214ms
cutoff:100000 threadCount:8 100times Time:8449ms
cutoff:100000 threadCount:16 100times Time:8479ms
cutoff:100000 threadCount:32 100times Time:8437ms
cutoff:100000 threadCount:64 100times Time:8438ms
cutoff:100000 threadCount:128 100times Time:8428ms
cutoff:100000 threadCount:256 100times Time:8539ms
cutoff:100000 threadCount:512 100times Time:8462ms
cutoff:100000 threadCount:1024 100times Time:8445ms
cutoff:100000 threadCount:2048 100times Time:8448ms
cutoff:100000 threadCount:4096 100times Time:8487ms
cutoff:100000 threadCount:8192 100times Time:8836ms
cutoff:100000 threadCount:16384 100times Time:8964ms
```

Below the output, there is an exception message:

```
Exception in thread "main" java.lang.IllegalArgumentException Create breakpoint
    at java.base/java.util.concurrent.ForkJoinPool.<init>(ForkJoinPool.java:2295)
    at java.base/java.util.concurrent.ForkJoinPool.<init>(ForkJoinPool.java:2165)
    at edu.neu.coe.info6205.sort.par.Main.main(Main.java:33)
```

The bottom status bar indicates the build completed successfully in 1 sec, 752 ms (4 minutes ago).

It comes the same results when changed the array.list to 1,000,000, 1,500,000, 2,500,000 and 3,000,000, as below.

IDEA screenshot showing a Java project with a bug in a parallel sorting algorithm. The code uses `ForkJoinPool` and `ParSort` to sort an array of 100,000 integers. The bug is an `IllegalArgumentException` thrown by `ForkJoinPool` due to an invalid thread count.

Code Snippet (Main.java):

```
public class Main {  
    public static void main(String[] args) {  
        processArgs(args);  
        System.out.println("Degree of parallelism: " + ForkJoinPool.getCommonPoolParallelism());  
        Random random = new Random();  
        int[] array = new int[100000];  
        ArrayList<Long> timeList = new ArrayList<>();  
        //for (int j = 1; j < 16; j++) { ParSort.cutoff = 10000 * j ;  
        ParSort.cutoff = 100000;  
    }  
}
```

Run Output (Top):

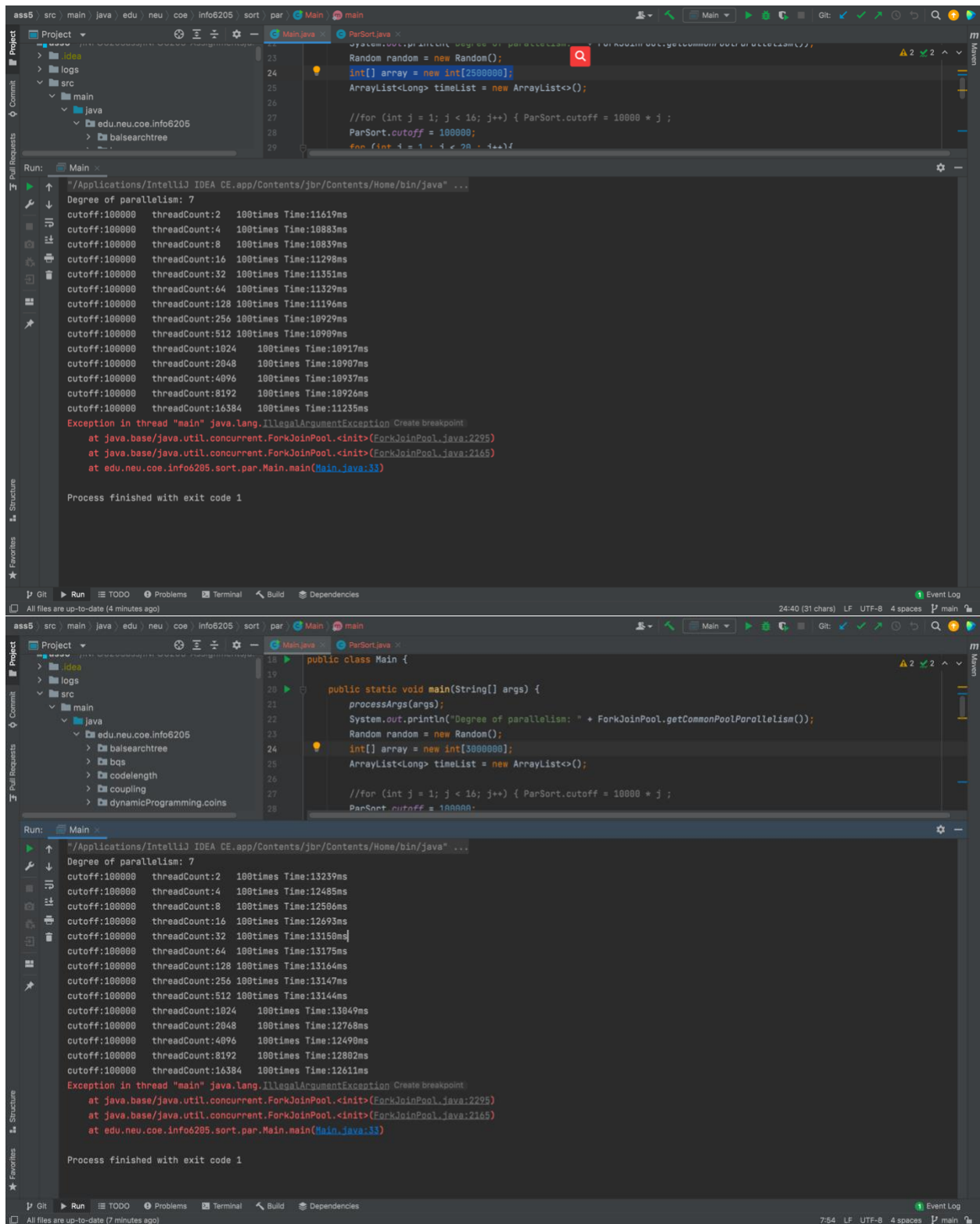
```
"/Applications/IntelliJ IDEA CE.app/Contents/jbr/Contents/Home/bin/java" ...  
Degree of parallelism: 7  
cutoff:100000 threadCount:2 100times Time:5101ms  
cutoff:100000 threadCount:4 100times Time:4498ms  
cutoff:100000 threadCount:8 100times Time:4579ms  
cutoff:100000 threadCount:16 100times Time:4546ms  
cutoff:100000 threadCount:32 100times Time:4535ms  
cutoff:100000 threadCount:64 100times Time:4525ms  
cutoff:100000 threadCount:128 100times Time:4516ms  
cutoff:100000 threadCount:256 100times Time:4732ms  
cutoff:100000 threadCount:512 100times Time:4584ms  
cutoff:100000 threadCount:1024 100times Time:4567ms  
cutoff:100000 threadCount:2048 100times Time:4522ms  
cutoff:100000 threadCount:4096 100times Time:4525ms  
cutoff:100000 threadCount:8192 100times Time:4529ms  
cutoff:100000 threadCount:16384 100times Time:4513ms  
Exception in thread "main" java.lang.IllegalArgumentException Create breakpoint  
at java.base/java.util.concurrent.ForkJoinPool.<init>(ForkJoinPool.java:2295)  
at java.base/java.util.concurrent.ForkJoinPool.<init>(ForkJoinPool.java:2165)  
at edu.neu.coe.info6205.sort.par.Main.main(Main.java:33)  
Process finished with exit code 1
```

Run Output (Bottom):

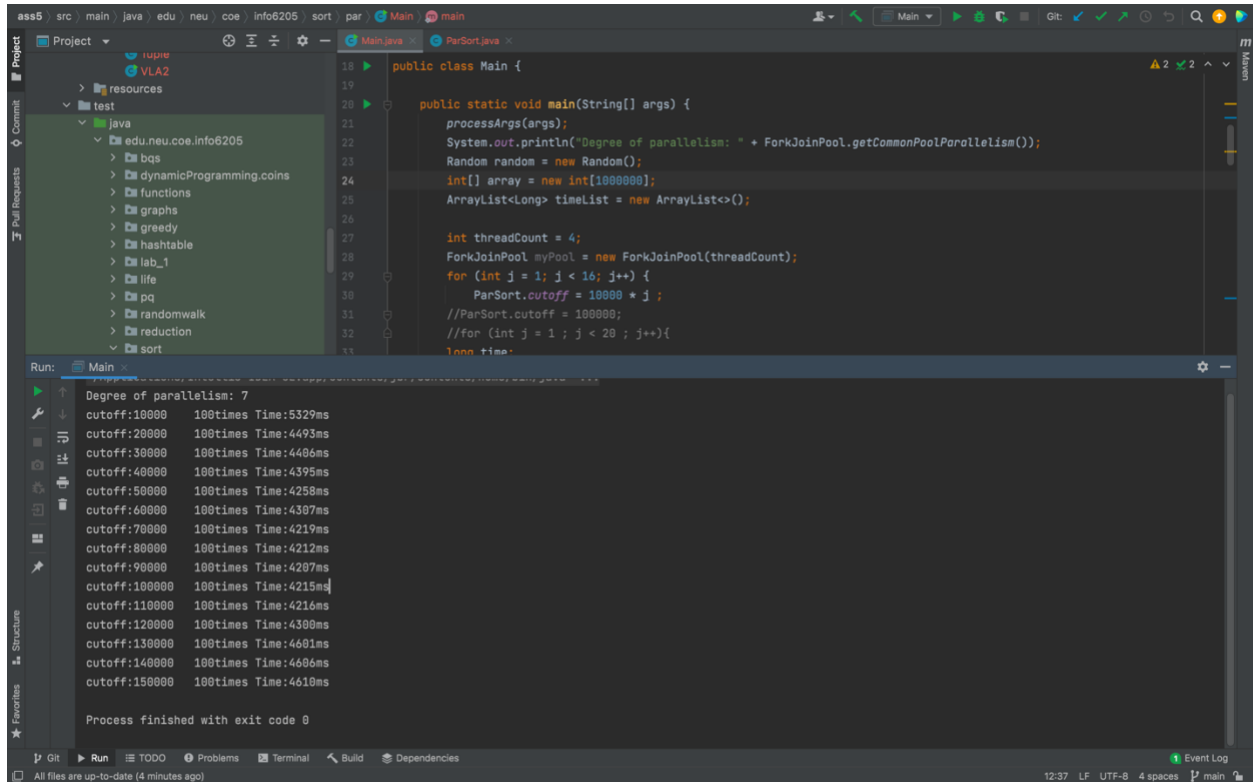
```
"/Applications/IntelliJ IDEA CE.app/Contents/jbr/Contents/Home/bin/java" ...  
Degree of parallelism: 7  
cutoff:100000 threadCount:2 100times Time:7147ms  
cutoff:100000 threadCount:4 100times Time:6574ms  
cutoff:100000 threadCount:8 100times Time:6661ms  
cutoff:100000 threadCount:16 100times Time:6690ms  
cutoff:100000 threadCount:32 100times Time:6713ms  
cutoff:100000 threadCount:64 100times Time:6708ms  
cutoff:100000 threadCount:128 100times Time:6697ms  
cutoff:100000 threadCount:256 100times Time:6675ms  
cutoff:100000 threadCount:512 100times Time:6698ms  
cutoff:100000 threadCount:1024 100times Time:6677ms  
cutoff:100000 threadCount:2048 100times Time:6660ms  
cutoff:100000 threadCount:4096 100times Time:6672ms  
cutoff:100000 threadCount:8192 100times Time:6676ms  
cutoff:100000 threadCount:16384 100times Time:6688ms  
Exception in thread "main" java.lang.IllegalArgumentException Create breakpoint  
at java.base/java.util.concurrent.ForkJoinPool.<init>(ForkJoinPool.java:2295)  
at java.base/java.util.concurrent.ForkJoinPool.<init>(ForkJoinPool.java:2165)  
at edu.neu.coe.info6205.sort.par.Main.main(Main.java:33)  
Process finished with exit code 1
```

Build Log:

```
Build completed successfully in 1 sec, 901 ms (6 minutes ago)
```



When the thread count is 4, for the array.length 1,000,000, 1,500,000, 2,000,000, 2,500,000 and 3,000,000, running cutoff from 10,000 to 150,000, the best value of cutoff is also around 100,000 as below.



The screenshot shows an IDE with a project named 'edu.neu.coe.info6205'. The code in 'Main.java' defines a 'Main' class with a 'main' method. The method sets 'threadCount' to 4, creates a 'ForkJoinPool', and runs a loop for 'j' from 1 to 16. Inside the loop, it sets 'ParSort.cutoff' to '10000 * j' and calls 'ParSort.sort' on an array of 10,000,000 integers. The 'Run' console shows the output of the program, including the 'Degree of parallelism' and a table of cutoff values and execution times.

```
public class Main {
    public static void main(String[] args) {
        processArgs(args);
        System.out.println("Degree of parallelism: " + ForkJoinPool.getCommonPoolParallelism());
        Random random = new Random();
        int[] array = new int[10000000];
        ArrayList<Long> timeList = new ArrayList<>();

        int threadCount = 4;
        ForkJoinPool myPool = new ForkJoinPool(threadCount);
        for (int j = 1; j < 16; j++) {
            ParSort.cutoff = 10000 * j;
            //ParSort.cutoff = 100000;
            //for (int j = 1; j < 20; j++){
            long time =
```

```
Degree of parallelism: 7
cutoff:10000 100times Time:5329ms
cutoff:20000 100times Time:4493ms
cutoff:30000 100times Time:4406ms
cutoff:40000 100times Time:4395ms
cutoff:50000 100times Time:4258ms
cutoff:60000 100times Time:4307ms
cutoff:70000 100times Time:4219ms
cutoff:80000 100times Time:4212ms
cutoff:90000 100times Time:4207ms
cutoff:100000 100times Time:4215ms
cutoff:110000 100times Time:4216ms
cutoff:120000 100times Time:4300ms
cutoff:130000 100times Time:4601ms
cutoff:140000 100times Time:4606ms
cutoff:150000 100times Time:4610ms

Process finished with exit code 0
```


IDE screenshot showing a Java project structure and a code editor. The project structure on the left includes a package `edu.neu.coe.info6205` with sub-packages like `bqs`, `dynamicProgramming.coins`, `functions`, `graphs`, `greedy`, `hashtable`, `lab_1`, `life`, `pq`, `randomwalk`, `reduction`, and `sort`.

The code editor displays the `Main.java` file with the following code:

```
public class Main {  
    public static void main(String[] args) {  
        processArgs(args);  
        System.out.println("Degree of parallelism: " + ForkJoinPool.getCommonPoolParallelism());  
        Random random = new Random();  
        int[] array = new int[1500000];  
        ArrayList<Long> timeList = new ArrayList<>();  
  
        int threadCount = 4;  
        ForkJoinPool myPool = new ForkJoinPool(threadCount);  
        for (int j = 1; j < 16; j++) {  
            ParSort.cutoff = 10000 * j;  
            //ParSort.cutoff = 100000;  
            //for (int j = 1; j < 20; j++){  
                long time;  
            }  
        }  
    }  
}
```

The Run console shows the output of the program:

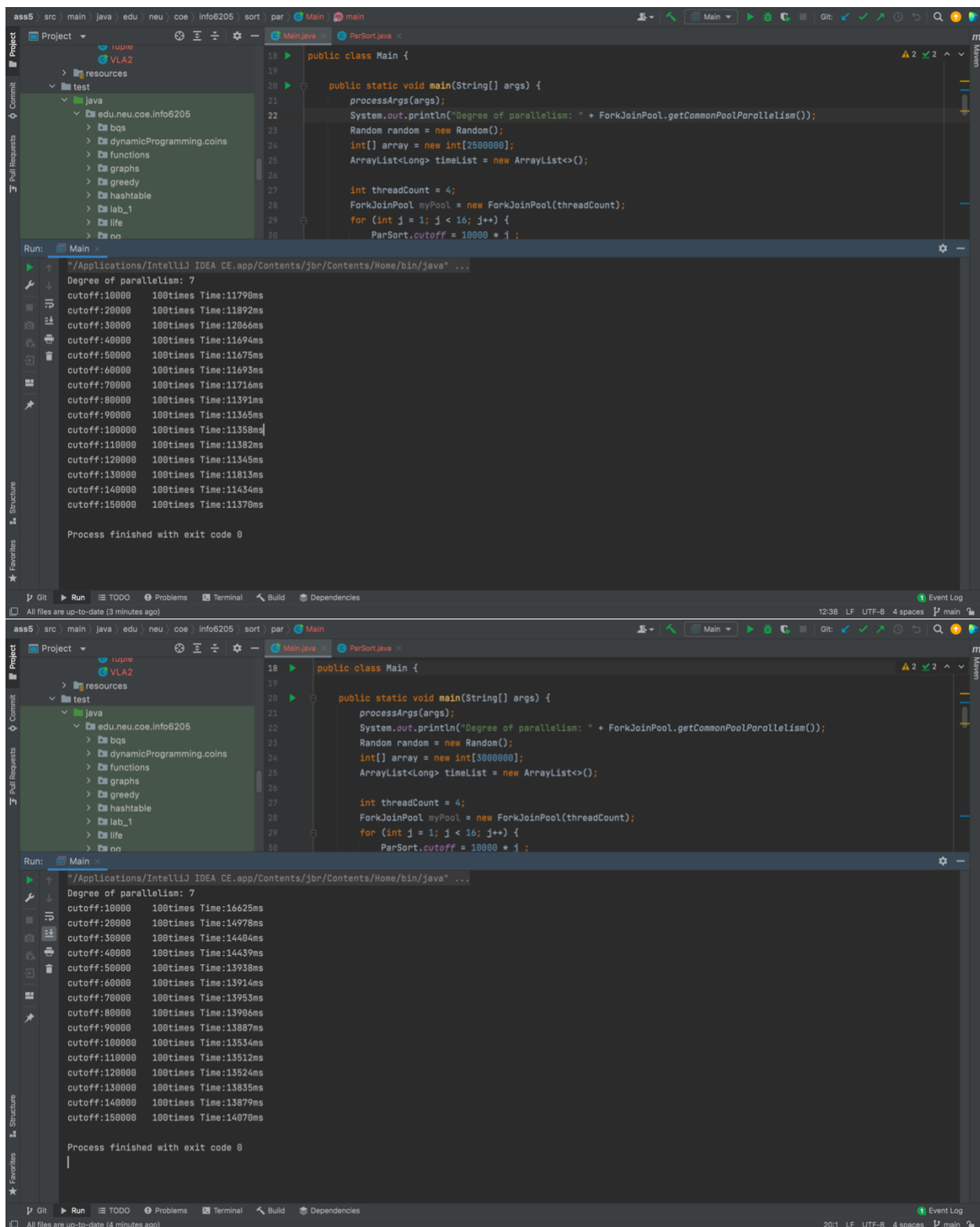
```
Degree of parallelism: 7  
cutoff:10000 100times Time:7576ms  
cutoff:20000 100times Time:6716ms  
cutoff:30000 100times Time:6424ms  
cutoff:40000 100times Time:6347ms  
cutoff:50000 100times Time:6287ms  
cutoff:60000 100times Time:6189ms  
cutoff:70000 100times Time:6192ms  
cutoff:80000 100times Time:6191ms  
cutoff:90000 100times Time:6201ms  
cutoff:100000 100times Time:6177ms  
cutoff:110000 100times Time:6275ms  
cutoff:120000 100times Time:6437ms  
cutoff:130000 100times Time:6420ms  
cutoff:140000 100times Time:6434ms  
cutoff:150000 100times Time:6401ms  
  
Process finished with exit code 0
```

IDE screenshot showing the same Java project structure and code editor. The code editor displays the `Main.java` file with the following code:

```
public class Main {  
    public static void main(String[] args) {  
        processArgs(args);  
        System.out.println("Degree of parallelism: " + ForkJoinPool.getCommonPoolParallelism());  
        Random random = new Random();  
        int[] array = new int[2000000];  
        ArrayList<Long> timeList = new ArrayList<>();  
  
        int threadCount = 4;  
        ForkJoinPool myPool = new ForkJoinPool(threadCount);  
        for (int j = 1; j < 16; j++) {  
            ParSort.cutoff = 10000 * j;  
            //ParSort.cutoff = 100000;  
            //for (int j = 1; j < 20; j++){  
                long time;  
            }  
        }  
    }  
}
```

The Run console shows the output of the program:

```
Degree of parallelism: 7  
cutoff:10000 100times Time:10012ms  
cutoff:20000 100times Time:9071ms  
cutoff:30000 100times Time:8904ms  
cutoff:40000 100times Time:8727ms  
cutoff:50000 100times Time:8708ms  
cutoff:60000 100times Time:8714ms  
cutoff:70000 100times Time:8599ms  
cutoff:80000 100times Time:8551ms  
cutoff:90000 100times Time:8558ms  
cutoff:100000 100times Time:8555ms  
cutoff:110000 100times Time:8593ms  
cutoff:120000 100times Time:8613ms  
cutoff:130000 100times Time:8587ms  
cutoff:140000 100times Time:8627ms  
cutoff:150000 100times Time:8575ms  
  
Process finished with exit code 0
```



2. Graphical Representation(Observations from experiments should be tabulated and analyzed by plotting graphs(usually in excel) to arrive on the relationship conclusion)

