Lab 1

Student information

- Full name: Qicheng Hu
- E-mail: qhu027@cs.ucr.edu
- UCR NetID: qhu027
- Student ID: X675102

Answers

- (Q1) What is the name of the created directory?

  The created directory was named 'qhu027_lab1'.

- (Q2) What do you see at the console output?

  The console displays 'Hello World!' as the following screenshot shows:

  ```
  (base) QichengdeMacBook-Pro:qhu027_lab1 qhu$ java -cp target/qhu027_lab1-1.0-SNA
  PSHOT.jar  edu.ucr.cs.cs167.qhu027.App
  Hello World!
  ```

- (Q3) What do you see at the output?

  Due to the reason that no input file is specified (no argument is set), the program throws an exception as the following screenshot shows:
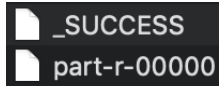
  ```
  Exception in thread "main" java.lang.ArrayIndexOutOfBoundsException: Index 0 out of bounds for length 0
      at edu.ucr.cs.cs167.qhu027.App.main(App.java:63)
  ```

- (Q4) What is the output that you see at the console?

  The console only displays the exit code of the program,

  ```
  Process finished with exit code 0
  ```

  but the output.txt folder is createed, which includes the following files:

In the part-r-00000 file, the following content is written :

```
but 1
cannot 3
crawl  1
do 1
fly,   1
forward    1
have   1
if 3
keep   1
moving 1
run    1
run,   1
then   3
to 1
walk   1
walk,  1
whatever   1
you    5
```

- (Q5) Does it run? Why or why not?

  If using java, it cannot run because of no hadoop class is found:

```
Exception in thread "main" java.lang.NoClassDefFoundError: org/apache/hadoop/con
f/Configuration
        at edu.ucr.cs.cs167.qhu027.App.main(App.java:55)
Caused by: java.lang.ClassNotFoundException: org.apache.hadoop.conf.Configuratio
n
        at java.net.URLClassLoader.findClass(URLClassLoader.java:382)
        at java.lang.ClassLoader.loadClass(ClassLoader.java:418)
        at sun.misc.Launcher$AppClassLoader.loadClass(Launcher.java:355)
        at java.lang.ClassLoader.loadClass(ClassLoader.java:351)
        ... 1 more
```

  If run using hadoop

```
hadoop jar target/qhu027_lab1-1.0-SNAPSHOT.jar  edu.ucr.cs.cs167.qhu027.App
input.txt output.txt
```

  then it runs and gives the correct result:

```
                Combine input records=28
                Combine output records=18
                Reduce input groups=18
                Reduce shuffle bytes=208
                Reduce input records=18
                Reduce output records=18
                Spilled Records=36
                Shuffled Maps =1
                Failed Shuffles=0
                Merged Map outputs=1
                GC time elapsed (ms)=5
                Total committed heap usage (bytes)=468713472
        Shuffle Errors
                BAD_ID=0
                CONNECTION=0
                IO_ERROR=0
                WRONG_LENGTH=0
                WRONG_MAP=0
                WRONG_REDUCE=0
        File Input Format Counters
                Bytes Read=140
        File Output Format Counters
                Bytes Written=142
```