CPEN522 Assignment 4

Ziyan Gong/94478161, Zhen Wang/98552169

1 Automated Test Generation

Monkey

```
Events injected: 10000
:Sending rotation degree=0, persist=false
:Dropped: keys=1 pointers=5 trackballs=0 flips=30 rotations=0
## Network stats: elapsed time=51308ms (Oms mobile, Oms wifi, 51308ms not connected)
// Monkey finished
```

Time: As the log shows, it takes 51308ms to generate 10000 random tests.

FindBug: We introduced an Arithmetric Bug in "MovieDetailsFragment.java". Monkey found the bug and here is the output report.

```
CRASH: com. esoxjem. movieguide (pid 9406)
| Short Msg: java. lang. ArithmeticException | divide by zero |
| Build Label: HLAWEL/EVA-ALOO/HWEVA:S. 0. O/HLAWELEVA-ALOO/540 (COO) : user/release-keys |
| Build Changelist: 540 (COO) |
| Build Time: 1545184313000 |
| java. lang. RuntimeException: Unable to start activity ComponentInfo (com. esoxjem. movieguide/com. esoxjem. movieguide. details. MovieDetailsActivity]: java. lang. ArithmeticException: divide by zero |
| ArithmeticExcepti
```

Evaluation: In general, monkey is easy to use for me. It is a command-line tool th at you can run on any emulator instance or on a device with the target APK install ed. Monkey is included in Android Studio, so you can launch it in the adb shell directly. This tool help us avoid some boring and repetitive testing operations.

Limitations: After several tests, I found some limitations of Monkey. Firstly, the e target testing object for Monkey must be apk. Secondly, the stream of user event is generated randomly. If a developer wants to reproduce the error, he must use the same pseudo-random seed. Thirdly, a tester need to unlock the screen before the etesting, otherwise Monkey couldn't enter the app correctly. Besides, when interacting with the system, the tools in toolbar are easily mistouched, which might information luence the testing, like turning off WiFi.

Improvements: If Monkey could take screenshot when detecting errors, that would be better for developers to understand abnormal scenarios.

Randoop

Time: The class list is set to include all the classes in JPacman shown as below.

```
org.jpacman.framework.factory.DefaultGameFactory
org.jpacman.framework.factory.FactoryException
org.jpacman.framework.factory.IGameFactory
org.jpacman.framework.factory.IGameFactory
org.jpacman.framework.factory.MapParser
org.jpacman.framework.controller.AbstractGhostMover
org.jpacman.framework.controller.IController
org.jpacman.framework.model.Board
org.jpacman.framework.model.Board
org.jpacman.framework.model.Food
org.jpacman.framework.model.Game
org.jpacman.framework.model.IGameInteractor
org.jpacman.framework.model.IGameInteractor
org.jpacman.framework.model.IPointInspector
org.jpacman.framework.model.Player
org.jpacman.framework.model.Player
org.jpacman.framework.model.Sprite
org.jpacman.framework.model.NointManager
org.jpacman.framework.model.Tile
org.jpacman.framework.model.Wall
org.jpacman.framework.model.Wall
org.jpacman.framework.ui.ButtonPanel
org.jpacman.framework.ui.Disposable
org.jpacman.framework.ui.PacmanInteraction
org.jpacman.framework.view.Animator
org.jpacman.framework.view.BoardView
org.jpacman.framework.view.BoardView
org.jpacman.framework.view.ImageLoader
```

By setting the time limit to 1 minute, it generates 29 error-revealing tests and 1 320 regression tests shown as below.

```
Average method execution time (normal termination): 0.0187
Average method execution time (exceptional termination): 0.089
Error-revealing test output:
Error-revealing test count: 29
Meximan Jubi-tests

Created file /Users/zhenwang/Desktop/cpen522/group11/jpacman-framework/src/test/java/org/jpacman/test/framework/ErrorTest0.java
Regression test output:
Regression test output:
Regression test count: 1320
Meximan Jubi-tests

A test code assertion failed during flaky-test filtering. Most likely,
you ran Randoop on a program with nondeterministic behavior. See section
'Most: RegressionTest4, Mothod test809, Line number: 715, Source line:

Class: RegressionTest4, Mothod test809, Line number: 715, Source line:

DuttonPanela.start();

Containing method:

if (debug)

Gystonic riormati('Moskine' "RegressionTest0.test809');

Grest
public void test809() throws Throwable {

if (debug)

Gystonic riormati('Moskine' "RegressionTest0.test809');

Grest
public void test809() throws Throwable {

if (abug)

Gystonic riormati('Moskine' "RegressionTest0.test809');

Grest
public void test809() throws Throwable {

if (abug)

Gystonic riormati('Moskine' "RegressionTest0.test809');

Grest
public void test809() throws Throwable {

if (abug)

Gystonic riormati('Moskine' "RegressionTest0.test809');

Grest
public void test809() throws Throwable {

if (abug)

Gystonic riormatic riormat
```

Coverage: The generated tests are saved as .java file with 29 tests for error-reve aling test file and 500 for one regression test file. They are put into the JPacma n project and the jar packages of JUnit and Hamcrest are put imported to the class path of eclipse.

By running the tests of error-revealing and regression test together, the coverage is as shown below.

	82.0 %	54,169	11,881	66,050
	82.6 %	51,406	10,798	62,204
	85.8 %	51,406	8,538	59,944
	0.0 %	0	1,712	1,712
	0.0 %	0	194	194
	0.0 %	0	161	161
	0.0 %	0	151	151
	0.0 %	0	42	42
	71.8 %	2,763	1,083	3,846
	68.9 %	1,018	460	1,478
-	70.4 %	718	302	1,020
1	30.2 %	60	139	199
	74.7 %	296	100	396
	89.1 %	671	82	753
		82.6 % 85.8 % 0.0 % 0.0 % 0.0 % 0.0 % 0.0 % 1 71.8 % 68.9 % 70.4 % 1 30.2 % 74.7 %	82.6 % 51,406 85.8 % 51,406 0.0 % 0 0.0 % 0 0.0 % 0 0.0 % 0 0.0 % 0 1 71.8 % 2,763 68.9 % 1,018 70.4 % 718 1 30.2 % 60 74.7 % 296	82.6 % 51,406 10,798 85.8 % 51,406 8,538 0.0 % 0 1,712 0.0 % 0 194 0.0 % 0 161 0.0 % 0 151 0.0 % 0 42 1 71.8 % 2,763 1,083 68.9 % 1,018 460 70.4 % 718 302 1 30.2 % 60 139 74.7 % 296 100

The coverage for the source code is 71.8% which is not bad but not very good, especially for the controller, the coverage rate is only 30.2% and for model which contribute the most codes, the coverage rate is only 68.9%. Also a lot of failures are found during the process which are all caused by nullPointerException.

FindBug: By running the error revealing test and the regression test separately, we can see the expected result that no pass for error revealing test and no error for the regression test.

To dig deeper into why the failures occur, the example is given to explain.



```
@Override
public Player makePlayer() {
    assert getGame() != null;
    Player p = new Player();
    getGame().addPlayer(p);
    return p;
}
```

Obviously, it is caused by the assertions in the source code.

It definitely can find errors with error-revealing test.

For the errors it found, the example is shown below to analyze the reason.

```
test01 (0.000 s
                                                   Failure Trace
test02 (0.000 s)
                                                   iava.lang.NullPointerException
test03 (0.000 s)
                                                    at org.jpacman.framework.model.Game.died(Game.java:199)
test04 (0.000 s)
                                                    at org.ipacman.test.framework.ErrorTest0.test02(ErrorTest0.iava:25)
test05 (0.000 s)
 @Test
 public void test02() throws Throwable {
      if (debug)
          System.out.format("%n%s%n", "ErrorTest0.test02");
      org.jpacman.framework.model.Game game0 = new org.jpacman.framework.model.Game();
      boolean boolean1 = game0.died();
 @Test
             S
             @Override
             public boolean died() {
                     return !getPlayer().isAlive();
             }
```

Test02 tests the game.died() method, but there is no player created and getPlayer () method response with a reference referring to no object. The reference points t o an address with java default null, and null do not have isAlive() method so the exception is raised and the failure is caused.

Yes, it is easy to use. The users just need to set the arguments of the command l ine properly according to the document and the error-revealing test and regression test will be generated.

Limitations:

It does have a decent coverage which is shown in the previous questions, but the c overage is not high enough compared to what we do to use manually created test suites in the previous assignments. By analyzing the coverage in the source code, the most obvious reason that leads to low coverage is that it does not cover all the p aths considering conditioning.

Also according to the previous question, the error-revealing test does find bugs (or errors). But under complicated implementation conditions provided in the source code, most of the nullPointerException will never occur in real runtime.

Improvements:

Improving the coverage should be possible. If it can integrate randoop with other systematic testing method properly or provide user-friendly methods for people to use it together with other popular systematic testing tools, the coverage may get improved.