

Program Repairing using Exception Types, Constraint Automata and Typestate

Draft 4.0

Wednesday 27th August, 2014

Abstract

Changes done

Runtime Exceptions are common types of exceptions which may lead to system crash which leads to shutdown or restart. For may critical application such scenario is unacceptable due to their nature which requires availability of the service. Program bugs which causes runtime exceptions often go unnoticed at the time of development as these exceptions are unchecked exceptions. The key issue is to guide the program through some exception suppression procedure which will leads the program to a consistent state hence improve the chance of surviving a fatal crash. Here we consider such programs for which restart is not an option.

In this paper, we present a novel technique to recover from unexpected runtime exceptions. We have used hybrid of two techniques for efficient detection of potential point of failure and patch it closest to that to minimize the damage. One technique uses type of runtime exception to apply appropriate patch. The other technique will provides typestate analysis technique which will detect typestate violations to apply the right patch.

General Terms Reliability, Languages

Keywords program repair, runtime exception, software patching, symbolic execution, static analysis, type-state

1. Introduction

Changes done

Exception handling attributes to the response of program during runtime to some exceptional condition encounter. Most of the time it changes normal flow of program. In many cases exception handling is natural part of software execution due to the nature of the software. An application which constantly accesses I/O which also includes share resources may throw exception if another application blocks it. Here in this paper we discuss and analyze java exceptions and produce repair patch based on that. Java supports two types of exceptions :

- **Checked exception** which requires explicit *throws* declaration at the method declaration or *try-catch* block by the developers. Such exceptions are handled carefully as they often involves accessing resources like network, database, file system, I/O etc.
- **Unchecked exception** which does not enforce similar handling mechanism as the former one. *java.lang.RuntimeException* and its subclasses and *java.lang.Error* are types of unchecked exceptions. *NullPointerException*, *ArrayIndexOutOfBoundsException*, *ArithmeticException* are examples of common java runtime exceptions.

Oracle official documentation says that “*Here’s the bottom line guideline: If a client can reasonably be expected to recover from an exception, make it a checked exception. If a client cannot do anything to recover from the exception, make it an unchecked exception*”. Unchecked exception, particularly runtime exceptions can be thrown from any point in the program making them quite unpredictable in nature. Due to this extensive testing phase is required to eliminate any bugs and solve corner cases. Yet many applications suffer unexpected runtime exception causing system crash which leads to shutdown or restart.

We find out many applications where system shutdown/restart is expensive due to their nature. Notable examples are air traffic control, auto pilot, life support system, smart power grids, telephone networks, robots like UAV and rovers deployed for surveillance, reconnaissance and knowledge acquisition in remote locations etc. These applications are real-time sensitive and there is no room for exception handling in such system. Sudden crash involves risk of human life, expensive equipments and critical services. Other example includes web applications which uses scripts to dynamically generate websites and interfaces as per customer preferences. Many E-commerce websites handles queries, access and process customer and shopping items data and commits large amount of transactions. Sudden system crash may result in loss of precious time and data which eventually may result in a frustrated customers move to other websites. Many time bad or malicious code leads to some vulnerability to critical applications and website which can be exploited by attack to orchestrate system crash. Thought these examples cover a large variety of applications, all of them point to some concern of *availability*.

Usually, developers tests their code in series of verifications which involves code review, static and dynamic analysis of the code, generate test cases to cover as much potential input .Yet may corner cases can be left overlooked which can cause runtime exceptions. Multi-threaded applications are also susceptible to erroneous thread interleaving. One such exception is *java.lang.IllegalMonitorStateException*, when a thread has attempted to wait on an object’s monitor or to notify other threads waiting on an object’s monitor without owning the specified monitor. Applications under adversarial situation should be considered

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CONF ’yy, Month d–d, 20yy, City, ST, Country.
Copyright © 20yy ACM 978-1-xxxx-xxxx-n/yy/mm...\$15.00.
<http://dx.doi.org/10.1145/nnnnnnnn.nnnnnnn>

where deliberate malicious input may cause it to fail. To recover from such situation, a mechanism is needed which can predict failure by doing invariant and symbolic analysis. Invariant analysis will detect particular variables outside legal/safe bound. Symbolic analysis will indicate to the potential point of failure.

In this paper we proposed two solution to suppress runtime example and ensure system survivability. The approach consists of four primary phases

- **Generate input data-set** : We index user input along with the global variables and method arguments of successful runs. The local variables are not indexed as they can be re-generated. These data-set is used as a reference to later executions which encounters runtime exceptions. Appropriate user input of previous successful run is chosen in terms of correlation coefficient.
- **Program slice for patching** : We perform static analysis prior to running the program to determine data dependencies of the variables. The analysis yields a dependency graph which is used to determine optimal slice to be used as patch. This patch is placed in catch block and executed with the values of previous successful run while the original code is wrapped in try block.
- **Determine type of exception and patching** : The characteristics of patching is dependent on the type of runtime exception encountered by the program. A piece of code may throws multiple types of exceptions and all of them are handled at the time of patching by instrumenting multiple catch blocks.
- **Use typestate for repairing** : Typestate analysis, sometimes called protocol analysis defines valid sequences of operations that can be typically modeled using Finite State Machine (FSM) where the states represent abstract state of the program and the symbols are certain method invocations to perform state transition. Typestates are capable of representing behavioral type refinements like Iterators, where *hasNext()* method should be called before the *next()* method call. Typestate analysis is widely used as a safety feature to ensure a certain sequence of operations maintains proper protocol or not. The documentations of the API used in the application will define the valid typestate for repairing.

The object of the patching is to repair the problem closest to it to minimize any collateral damage to other parts of the applications hence minimizing the chance of unintentional data loss/corruption.

2. Motivation and Challenges

To Do

3. Problem Formulation

This part is incomplete, I am now writing the strategy part

We formulate the problem in following way

3.1 Runtime Exceptions

We can visualize all runtime exceptions as finite state machine (FSM). When a program violates such sequence, it throws runtime exception. In Figure 1, array index out of bound (java.lang.ArrayIndexOutOfBoundsException) exception is described as a FSM. Here, a program will be in safe bound as long as the $array_index \geq 0$ or $array_index \leq max_array_size - 1$

4. Repairing Strategy:Exception Type

Please review this section

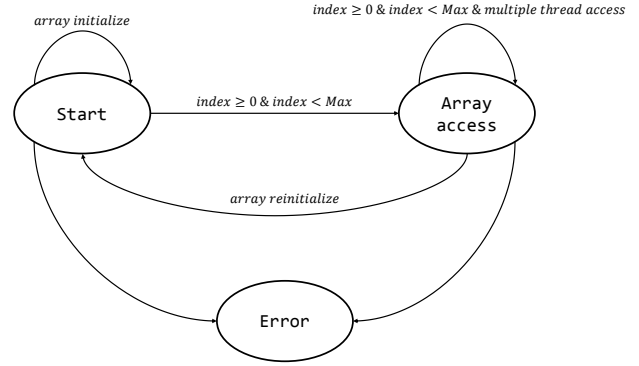


Figure 1. array index out of bound formulated as FSM

Listing 1. Java code which may throws runtime exceptions

```

1
2 public class TestClass
3 {
4     private int[] arr1;
5     private int[] arr2;
6     private int[] arr3;
7
8     public TestClass(int[] arr1, int[] arr2, int[]
9         arr3)
10    {
11        this.arr1 = arr1;
12        this.arr2 = arr2;
13        this.arr3 = arr3;
14    }
15
16    public int[] fun(int a, int b, int c, int d)
17    {
18        int temp0 = a + b;
19        int temp1 = c * d;
20        int temp2 = temp0 - temp1;
21        //array index out of bound, negative index
22        int temp3 = this.arr1[temp0];
23        //array index out of bound, negative index
24        int temp4 = this.arr2[temp1];
25        //array index out of bound, negative index
26        int temp5 = this.arr3[temp3];
27        int temp6 = temp4 + temp5;
28        int temp7 = temp6 - temp3;
29        //array index out of bound, negative
30        //index, divide by zero
31        this.arr1[temp6] = temp7/(d-a);
32        //array index out of bound, negative
33        //index, divide by zero
34        this.arr2[temp7] = temp7/temp4;
35        if(arr2[temp1] != arr3[temp7])
36            return arr1;
37        else
38            return null;
39    }
40 }
41
42 public class MainClass
43 {
44     public void main(String[] a)
45     {
46         int[] arr1 = {1,2,3,4};
47         int[] arr2 = {1,2,3,4};
48         int[] arr3 = {1,2,3,4};
49         TestClass TC = new TestClass(arr1, arr2,
50             arr3);
51         int[] res = TC.fun(2,4,3,4);
52         //Null pointer exception
53         System.out.print("Result : "+res[2]);
54     }
55 }

```

48 }
49 }

In the Example 1, we have given a piece of java code which shows multiple lines can throw several runtime exceptions. In this example we consider three very common runtime exceptions: `NullPointerException`, `ArrayIndexOutOfBoundsException`, `NegativeIndexException`, `ArithmeticException` (i.e. divide-by-zero). In rest of this section, this particular example will be used to demonstrate the repairing strategy.

4.1 Symbolic Analysis

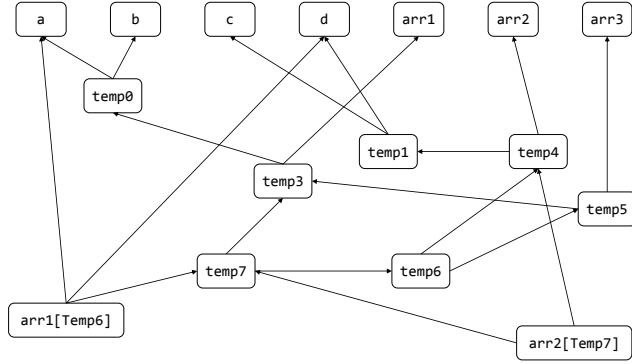


Figure 2. Data dependency graph of the variables in Example 1

We have done several static analysis a priori over the Java source code to discover :

1. Critical section of the code which are not eligible for patching. Eg. banking or any financial transaction which should be crashed in case of exception as suboptimal solution due to patching will led it to inconsistent state.
2. Symbolic analysis of the program to discover potential points of failure and mark them.
3. Build data dependency graph which will be used to generate appropriate code slice to be used as patch. In Figure 2, the data dependency graph of the example code 1 is presented.
4. The symbolic analysis will also reveal which kind of exception is likely to happened at the time of execution. This information is necessary at the time of instrumenting the patch as it will determine the catch block.

4.2 Data set for Successful Program Runs

Here we will store all the traces of successful program runs.

Global variables and parameters						
a	b	c	d	arr1	arr2	arr3
<snapshot>

Successful runs

Figure 3. Indexed global variables and method arguments successful runs

Figure 3 shows such indexed traces of all the global variables and method arguments. We store the snapshots of these objects. We won't store local variables as they can always be regenerated. As it is required to capture the snapshot of all these variable, we made deep clone of all of these objects and variables.

4.3 Matrices

Please review this section.

4.4 Instrumenting Patching

We have used Soot framework which is a Java byte code manipulator to instrument patch. The patching technique is divided into two phases

4.4.1 Determine Exception Type

At the time of execution, the exception may happened due to some specific values of some variables. We will catch the exception. Here the type of runtime exception is `java.lang.ArrayIndexOutOfBoundsException`. This will be used to produce the try-catch block.

4.4.2 Determine Optimal Code Slice

The optimal code slice will be determined from the data dependency graph which was rendered at the time of static analysis mentioned in Section 4.1. In the Listing 2, the example code snippet shows such code slice inside the catch block. As the error occurred at the line `int temp5 = this.arr3[temp3]`; the statements which produces the temp3 and the statement which also involves temp3 or any other variables derived from temp3, would be included in the catch block for re-execution with the valued of the same from the data table of previous successful runs.

Listing 2. patching code slice based on exception type

```

1 public class TestClass
2 {
3     private int[] arr1;
4     private int[] arr2;
5     private int[] arr3;
6
7     public TestClass(int[] arr1, int[] arr2, int[]
8         arr3)
9     {
10         this.arr1 = arr1;
11         this.arr2 = arr2;
12         this.arr3 = arr3;
13     }
14     public int[] fun(int a, int b, int c, int d)
15     {
16         try
17         {
18             int temp0 = a + b;
19             int temp1 = c * d;
20             int temp2 = temp0 - temp1;
21             int temp3 = this.arr1[temp0];
22             int temp4 = this.arr2[temp1];
23             //IndexOutOfBoundsException as temp3 = 20
24             int temp5 = this.arr3[temp3];
25             int temp6 = temp4 + temp5;
26             int temp7 = temp6 - temp3;
27             this.arr1[temp6] = temp7/(d-a);
28             this.arr2[temp7] = temp7/temp4;
29         }
30         catch (IndexOutOfBoundsException indEx)
31         {
32             int temp0 = a + b;
33             int temp1 = c * d;
34             int temp2 = temp0 - temp1;
35             int temp3 = this.arr1[temp0];
36             //Bellow line is not part of the patch as
37             //temp1 and temp3 are not related to temp3
38             //for which the exception occurred.
39             //int temp4 = this.arr2[temp1];
40             int temp5 = this.arr3[temp3];

```

```

41         }
42         if(arr2[temp1] != arr3[temp7])
43             return arr1;
44         else
45             return null;
46     }
47 }
48 public class MainClass
49 {
50     public void main(String[] a)
51     {
52         int[] arr1 = {20,21,22,23};
53         int[] arr2 = {1,2,3,4};
54         int[] arr3 = {10,11,12,13};
55         TestClass TC = new TestClass(arr1, arr2,
56                                     arr3);
57         int[] res = TC.fun(2,4,3,2);
58         System.out.print("Result : "+res[2]);
59     }
60 }

```

4.5 Variable Tracking and Monitoring

I have added standard taint analysis technique here as an example. We can change it later

Here we used taint analysis technique to tag variables and objects of our interest to monitor them. This steps are necessary as the values of the variables used during the instrumentation may cause further runtime exceptions. We used bit-vector which is an efficient technique to taint a object/variable. It requires maintain a single dimension byte array where each bit correspond to a single object/variable of our interest. The bit values will be flipped when it is required to taint (1) or untaint (1) an object/variable. We will only monitor these entities until all of them flushed from the program and the entire program reached to a stable state.

5. Repairing Strategy : Constraint Automata

5.1 General Structure

Constraint automata is a formalism to describe the behavior and possible data flow in coordination models. Mostly used for model checking. We have used it for the purpose of program repairing technique. Here we define the finite state automata as follows :

$$(Q, \Sigma, \delta, q_0, F)$$

- Q : set of state where $|Q| = 2$, *legal state*(init) and *illegal state* (error).
- Σ : symbols, invariants based on exception type.
- δ : transition function. *init* \rightarrow *init* is safe transition and *init* \rightarrow *error* is the invariant violation.
- q_0 : starting state, here $q_0 = \text{init}$.
- F : end state, here it same as q_0 .



Figure 4. Constraint automata general model

According to the Figure 4, the repairing mechanism will only trigger when we have a transition from init state to error state due to invariant violation.

5.2 Patching Techniques

The patching technique is based on the exception type.

5.2.1 Array index out of bound exception

Array index out of bound exception happen when one tries to access the array with a index which is more than the size of the array or less than zero i.e. with some negative value. We did the patching based on these two scenario. When the index is more than the array size, we patch it by assigning `array.length - 1`.

Listing 3. array index out of bound patching

```

1 void foo()
2 {
3     int []arr = {1,2,3,4};
4     int index = 10;
5     int y = 0;
6     try
7     {
8         //original code
9         y = arr[index];
10    }
11    //patching instrumentation
12    catch(IndexOutOfBoundsException ex)
13    {
14        if(index > arr.length)
15            y = arr[arr.length - 1];
16        else
17            y = a[0];
18    }
19 }

```

5.2.2 Negative Array Size Exception

Negative array size exception occurs when one tries to create a array with a negative size. The patching is done based on data flow analysis. Suitable index size is determined by looking at the successive statement dependent on the array. To take a safe bound, we took maximum index size and set as the array size in the new array statement.

Listing 4. arr index out of bound patching

```

1 void foo()
2 {
3     int []arr = {1,2,3,4};
4     int index = 10;
5     int y = 0;
6     try
7     {
8         //original code
9         y = arr[index];
10    }
11    //patching instrumentation
12    catch(IndexOutOfBoundsException ex)
13    {
14        if(index > arr.length)
15            y = arr[arr.length - 1];
16        else
17            y = a[0];
18    }
19 }

```

5.2.3 Arithmetic Exception : Division-by-zero Exception

Division by zero causes arithmetic exception. There are two different cases which were considered here.

- **Case I** : The denominator is going to the taint sink but the left hand side is not going to any taint sink. Here we will not manipulate the denominator as we are not manipulating any variable which are going to any taint sink.
- **Case II** : The denominator and the left hand side, both are not going to any taint sink. So they are safe to patch.

Listing 5. arithmetic exception : division-by-zero patching

```
1 void foo()
2 {
3     int a = 10;
4     int b = 0;
5     int y;
6     try
7     {
8         //original code
9         y = a/b;
10    }
11    //patching instrumentation
12    catch(ArithmeticException ex)
13    {
14        //case I
15        if(taintSink(b))
16            y = 0;
17        //case II
18        else
19        {
20            b = 1;
21            y = a/b;
22        }
23    }
24 }
```

6. Benchmark Results

7. Related Work

8. Conclusion and Future Works

Acknowledgments