# An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

April 26, 2015

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

Project Overviev

Empirical Study EXAM

Empirical Stud

easibility=

Using CodeCover Parsing Output XML Representation Parsing XML

Conclusion

#### Table of Contents

- 1. Motivation
- 2. Project Overview
  - 2.1 Automatic Fault Localization
  - 2.2 Empirical Study
  - 2.3 CodeCover
- 3. Feasibility
  - 3.1 Using CodeCover
  - 3.2 Parsing Output
- 4. Conclusion

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

Project Overview

Empirical Study EXAM

Empirical Stud

Feasibility

Using CodeCover Parsing Output XML Representation

· a a altitude a a

\_ .

#### Motivation

▶ Debugging is complex and difficult

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

Project Overview

Empirical Study EXAM

Empirical Stud

easibility

Using CodeCover
Parsing Output
XML Representation

onclusion

#### Motivation

- Debugging is complex and difficult
- ► Fault localization is the most expensive

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

Project Overview

Empirical Study
EXAM

NCP Empirical Stud

CodeCover

easibility=

Jsing CodeCover
Parsing Output
XML Representation

onclusion

#### Motivation

- Debugging is complex and difficult
- ► Fault localization is the most expensive
- Current techniques insufficient

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Motivation

#### Automatic Fault Localization

Uses per-test coverage analysis

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

Duniant Our

AFL Empirical Study EXAM

Empirical Stud CodeCover

easibility

Using CodeCover
Parsing Output
XML Representation
Parsing XMI

onclusion

#### Automatic Fault Localization

- Uses per-test coverage analysis
- Ranks statements by suspiciousness

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

Project Overview

AFL Empirical Study EXAM

Empirical Stud

easibility

Jsing CodeCov

arsing Output XML Representatio Parsing XMI

onclusion

#### Automatic Fault Localization

- Uses per-test coverage analysis
- Ranks statements by suspiciousness
- Variety of risk evaluation functions

$$suspiciousness(e) = 1 - \frac{\frac{failed(e)}{totalfailed}}{\frac{passed(e)}{totalpassed} + \frac{failed(e)}{totalfailee}}$$

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

Project Overview

Empirical Study

NCP Empirical Stud

CodeCover

Feasibi

Using CodeCover
Parsing Output

XML Representation

Conclusion

#### **EXAM Score**

▶ One method for comparing risk evaluation functions

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

roiect Overview

AFL Empirical Study EXAM

NCP Empirical Study

easibility

Using CodeCover
Parsing Output
XML Representation
Parsing XMI

onclusion

#### **EXAM Score**

- ▶ One method for comparing risk evaluation functions
- Ranking of faulty statement in suspiciousness

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

Project Overview

AFL Empirical Study

NCP

Empirical Stud CodeCover

easibility

Jsing CodeCover
Parsing Output

XML Representation

onclusion

#### **FXAM Score**

- ▶ One method for comparing risk evaluation functions
- Ranking of faulty statement in suspiciousness
- Percentage of statements not needed to consider

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

**EXAM** 

#### **EXAM Score**

- ▶ One method for comparing risk evaluation functions
- Ranking of faulty statement in suspiciousness
- Percentage of statements not needed to consider
- ► Higher-is-better metric

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

Project Overview

AFL Empirical Study EXAM

NCP

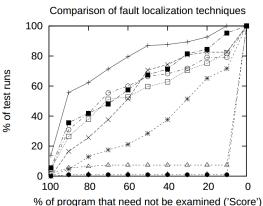
Empirical Study CodeCover

easibility

Jsing CodeCover Parsing Output XML Representation

Conclusion

#### **EXAM Score**





An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

\_

AFL

EXAM

Empirical Stud

CodeCover

Feasibilit

Jsing CodeCover
Parsing Output

XML Representation

Conclusion

 Alternate method for comparing risk evaluation functions An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

Project Overview

Empirical Study EXAM NCP

CodeCover

easibility

Jsing CodeCover
Parsing Output
XML Representation
Parsing XMI

onclusion

- Alternate method for comparing risk evaluation functions
- ► Risk evaluation function becomes heuristic for GenProg

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

Project Overview

Empirical Study
EXAM
NCP

Empirical Stu CodeCover

easibility

sing CodeCover arsing Output KML Representation

onclusion

- Alternate method for comparing risk evaluation functions
- ▶ Risk evaluation function becomes heuristic for GenProg
- Number of patches before correct patch

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

. . .

NFL Impirical Study

NCP Empirical Study

CodeCover

Feasibility

Jsing CodeCover Parsing Output XML Representation

onclusion

- Alternate method for comparing risk evaluation functions
- ► Risk evaluation function becomes heuristic for GenProg
- Number of patches before correct patch
- Lower-is-better metric

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

Project Overviev

Empirical Study
EXAM
NCP

Empirical Stu CodeCover

Feasibil

Jsing CodeCover Parsing Output XML Representation

onclusion

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener



Motivation



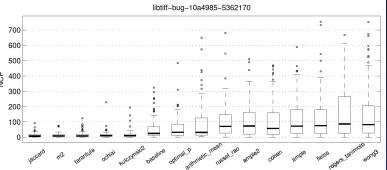
NCP Empirical Stu

CodeCover

#### Feasibil

Using CodeCover
Parsing Output
XML Representation

Conclusion



▶ Use EXAM score to compare risk evaluation functions

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

roject Overview

AFL Empirical Study EXAM

NCP Empirical Study

ancibility

Jsing CodeCover Parsing Output XML Representation

onclusion

- ▶ Use EXAM score to compare risk evaluation functions
- Select from existing studies

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

Project Overview

AFL Empirical Study

Empirical Study

easibility

Jsing CodeCover Parsing Output XML Representation Parsing XMI

onclusion

- ▶ Use EXAM score to compare risk evaluation functions
- Select from existing studies
- ► GenProg(NCP) comparison

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

roject Overvie

AFL Empirical Study

EXAM NCP

Empirical Study CodeCover

Feasibility

Jsing CodeCover
Parsing Output

XML Representation

Conclusion

- ▶ Use EXAM score to compare risk evaluation functions
- Select from existing studies
- ► GenProg(NCP) comparison
- ► Theoretical comparison

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

Project Overview

AFL Empirical Study

NCP Empirical Study

CodeCover

Feasibility

Using CodeCover
Parsing Output

XML Representation

Conclusion

► Eclipse-compatible coverage analysis tool

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

Project Overview

AFL Empirical Study

Empirical Stud

Eggeibility

easibility

Using CodeCover
Parsing Output

XML Representation

Parsing XMI

onclusion

- ► Eclipse-compatible coverage analysis tool
- Uses existing JUnit test suites

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

roject Overvie

AFL Empirical Study

NCP Empirical Stud

CodeCover

easibility

Using CodeCover
Parsing Output
XML Representation

onclusion

- ► Eclipse-compatible coverage analysis tool
- Uses existing JUnit test suites
- ► Generates coverage information for each test method

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

Project Overvice

AFL Empirical Study

NCP

CodeCover

easibility

sing CodeCover arsing Output KML Representation

onclusion

- ► Eclipse-compatible coverage analysis tool
- Uses existing JUnit test suites
- ▶ Generates coverage information for each test method
- ► Stores output in readable format (XML)

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

AFL Empirical Study

EXAM NCP

Empirical St CodeCover

Feasibility

Jsing CodeCover Parsing Output XML Representation

Conclusion

## Using CodeCover

► Tested with a simple existing system with a JUnit test suite

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

Project Overview

AFL Empirical Study FXAM

Empirical Study

easibility

Using CodeCover
Parsing Output
XML Representation
Parsing XMI

onclusion

## Using CodeCover

- ► Tested with a simple existing system with a JUnit test suite
- Successfully produced per-test coverage

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Lontents

Motivation

Project Overview

AFL Empirical Study FXAM

Empirical Stud

CodeCover

easibility

Using CodeCover Parsing Output XML Representation

onclusion

## Using CodeCover

- ► Tested with a simple existing system with a JUnit test suite
- Successfully produced per-test coverage
- Produced XML representation of coverage data

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

Project Overview

AFL Empirical Study

NCP Empirical Stud

CodeCover

Feasibility

Using CodeCover Parsing Output XML Representation

onclusion

## Coverage Highlighting

```
public static String listsToString(ArravList<ArravList<Object>> list)
           String ret = "":
           int listSize = list.size():
          for (int i = 0; i < listSize; i++)
                    += "L" + (i + 1) + ": \t" + list.get(i).
                      < listSize)
           return ret;
Problems @ Javadoc Declaration Console Problems
st Session Container: Listswap Oct 27, 2014 5:34:19 PM
dame
    edu.allegheny.test.ListSwapGeneratorTest:testMixed
    edu.alleghenv.test.ListSwapGeneratorTest:testListEmpty
    edu.allegheny.test.ListSwapGeneratorTest:testDouble
```

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

\_ . \_ \_

Project Overv

Empirical Study

Empirical Stud

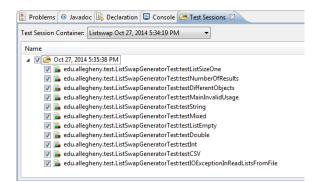
Feasibility

Feasibility

Using CodeCover Parsing Output XML Representation

Conclusion

### Per-test Coverage



An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

. . . .

AFL

Empirical Study EXAM

Empirical Stud

Feasibility

Using CodeCover Parsing Output

> XML Represent Parsing XML

. .

Conclusion

Contains several pieces of information:

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

. . . .

AFL Empirical Study

NCP Empirical Stud

CodeCover

easibility

Using CodeCover
Parsing Output
XML Representation
Parsing XMI

onclusion

- Contains several pieces of information:
  - Complete source code
  - Statement definitions
  - List of statements covered by each test method for each file under test

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

AFL Empirical Study EXAM

Empirical Stud

easibility

Using CodeCover
Parsing Output

XML Representation
Parsing XMI

Conclusion

- An Eclipse-based Integrated and Automated Fault Localization System
- Tristan Challener

- XML Representation

- Contains several pieces of information:
  - Complete source code
  - Statement definitions
  - List of statements covered by each test method for each file under test

<BasicStmnt CovItemId="S4" CovItemPrefix="edu.allegheny.listswap.ListSwapGenerator</p> <Tootist>

<Toc FndOffset="777" SrcFileId="1" StartOffset="735"/>

</LocList>

</BasicStmnt>

```
<TestCase Comment="" Date="1414365770750"</pre>
    Name="edu.allegheny.test.ListSwapGeneratorTest:testString">
<CovList>
<CovPrefix CovItemPrefix="edu.allegheny.listswap.ListSwapGenerator.jaya">
<Cov CovItemId="B1" Value="5"/>
<Cov CovItemId="B2" Value="1"/>
<Cov CovItemId="I.1-2" Value="1"/>
<Cov CovItemId="L2-0" Value="1"/>
<Cov CovItemId="L2-1" Value="1"/>
<Cov CovItemId="L2-2" Value="2"/>
<Cov CovItemId="L3-2" Value="1"/>
<Cov CovItemId="S1" Value="1"/>
<Cov CovItemId="S10" Value="1"/>
<Cov CovItemId="S11" Value="1"/>
<Cov CovItemId="S12" Value="1"/>
<Cov CovItemId="S13" Value="6"/>
<Cov CovItemId="S14" Value="6"/>
<Cov CovItemId="S15" Value="5"/>
<Cov CovItemId="S16" Value="1"/>
<Cov CovItemId="S2" Value="1"/>
<Cov CovItemId="S3" Value="4"/>
```

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

.........

Project Overview AFL

> mpirical Study EXAM

Empirical Stu

Feasibility

Using CodeCover
Parsing Output
XML Representation
Parsing XML

Conclusion

Document Object Model(DOM) parsing

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

Project Overview

Empirical Study
EXAM

Empirical Stud

easibility

Using CodeCover
Parsing Output
XML Representation
Parsing XML

. .

onclusion

- ► Document Object Model(DOM) parsing
- ► Translate entire XML document into Java tree structure

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

Project Overview

AFL Empirical Study EXAM

Empirical Study

CodeCover

easibility

sing CodeCover arsing Output (ML Representation

Parsing XML

Conclusion

- Document Object Model(DOM) parsing
- Translate entire XML document into Java tree structure
- ► Allows ease of handling data once built

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

Project Overview

Empirical Study EXAM

Empirical Study

CodeCover

Feasibility

Using CodeCover
Parsing Output
XML Representation
Parsing XML

.........

Conclusion

ELEM: nodeName="BasicStmnt"

ATTR: nodeName="ftext" nodeValue="S4"

TEXT: nodeName="ftext" nodeValue="s4"

ATTR: nodeName="ftext" nodeValue="edu.allegheny.listswap.ListSwapGenerator

TEXT: nodeName="ftext" nodeValue="edu.allegheny.listswap.ListSwapGenerator.jaya"

ATTR: nodeName="ftext" nodeValue="22"

TEXT: nodeName="ftext" nodeValue="22"

TEXT: nodeName="ftext" nodeValue=[WS]

ELEM: nodeName="ftext" nodeValue=[WS]

ELEM: nodeName="ftext" nodeValue="777"

TEXT: nodeName="LoCist"

TEXT: nodeName="ftext" nodeValue="777"

ATTR: nodeName="ftext" nodeValue="777"

ATTR: nodeName="ftext" nodeValue="1"

TEXT: nodeName="ftext" nodeValue="1"

TEXT: nodeName="ftext" nodeValue="1"

ATTR: nodeName="StartOffset" nodeValue="735"
TEXT: nodeName="#text" nodeValue="735"

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

......

Project Overview AFL

Empirical Study EXAM NCP

Empirical Study

Feasibility

Jsing CodeCover
Parsing Output
XML Representation

Parsing XML

Conclusion

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

ontents

Motivation

Project Overvie

AFL Empirical Stud FXAM

NCP Empirical Stud

CodeCover

Feasibility

reasibility

Parsing Output

XML Representation

Parsing XML

Conclusion

Conclusion

References

```
"I3-0" -0" P
```

TEXT: nodeName="#text" nodeValue="edu.allegheny.listswap.ListSwapGenerator.java"

TEXT: nodeName="#text" nodeValue=[WS]

ELEM: nodeName="CovItemId" nodeValue="L1-0"

ATTR: nodeName="#text" nodeValue="11-0"

ATTR: nodeName="#text" nodeValue="1"

TEXT: nodeName="#text" nodeValue="1"

TEXT: nodeName="#text" nodeValue=[WS]

ELEM: nodeName="Gov"

ATTR: nodeName="CovItemId" nodeValue="L3-0"

TEXT: nodeName="TovItemId" nodeValue="L3-0"

ATTR: nodeName="fext" nodeValue="1"

TEXT: nodeName="#text" nodeValue="1"

TEXT: nodeName="#text" nodeValue="1"

TEXT: nodeName="#text" nodeValue="1"

TEXT: nodeName="#text" nodeValue="I"

TEXT: nodeName="#text" nodeValue="I"

TEXT: nodeName="#text" nodeValue=[WS]

ATTR: nodeName="CovItemPrefix" nodeValue="edu.alleghenv.listswap.ListSwapGenerator

#### Conclusion

Create an Eclipse plugin that uses existing tools

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

Project Overview

AFL Empirical Study

NCP Empirical Stud

Eggeibility

Jsing CodeCover Parsing Output

Parsing XML

Conclusion

#### Conclusion

- Create an Eclipse plugin that uses existing tools
- Combine existing tools in a single fault localization system

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Contents

Motivation

Duniant Oursian

AFL Empirical Study

Empirical Study

---:

Jsing CodeCover
Parsing Output

XML Representation

Conclusion

#### Conclusion

- Create an Eclipse plugin that uses existing tools
- ► Combine existing tools in a single fault localization system
- Ambitious, but feasible

An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

Conclusion

#### References



James A. Jones and Mary Jean Harrold.

Empirical evaluation of the tarantula automatic fault-localization technique.

In Proceedings of the 20th IEEE/ACM International Conference on Automated Software Engineering, ASE '05, pages 273–282, New York, NY, USA, 2005. ACM.



Yuhua Qi, Xiaoguang Mao, Yan Lei, and Chengsong Wang.

Using automated program repair for evaluating the effectiveness of fault localization techniques. In Proceedings of the 2013 International Symposium on Software Testing and Analysis, ISSTA 2013, pages 191–201, New York, NY, USA, 2013. ACM.



Xiaoyuan Xie, Tsong Yueh Chen, Fei-Ching Kuo, and Baowen Xu.

A theoretical analysis of the risk evaluation formulas for spectrum-based fault localization. ACM Trans. Softw. Eng. Methodol., 22(4):31:1–31:40, October 2013. An Eclipse-based Integrated and Automated Fault Localization System

Tristan Challener

contents

Motivation

Project Overview

AFL Empirical Study EXAM

Empirical Stu

easibility

Using CodeCover Parsing Output XML Representation

Lonclusion