On Sufficiency of Mutants

Akbar Siami Namin and James H. Andrews
Department of Computer Science
The University of Western Ontario, Canada.

Doctoral Symposium ICSE 2007 - Minneapolis, USA

May 21, 2007

Outline

Motivation

Challenge

Problem Statement

Possible Solution

Our Approach

Preliminary Result

Conclusion

Future Work

Outline
Motivation
Motivation
Needs for experimental study
Challenge
Problem Statement
Possible Solution
Our Approach
Preliminary Result
Conclusion
Future Work

It is about the cost

Motivation

Outline

Motivation

- Motivation
- Needs for experimental study

Challenge

Problem Statement

Possible Solution

Our Approach

Preliminary Result

Conclusion

Future Work

- Current testing techniques are still expensive
- We need more effective and less expensive testing techniques
- How to assess the efficiency and effectiveness of the new techniques?
  - Efficiency: relatively easy to assess by defining some metrics (e.g., time, cost, used resources, etc.)
  - Effectiveness: Not easy to address!!!

Needs for experimental study

Outline

Motivation

Motivation

Needs for experimental study

Challenge

Problem Statement

Possible Solution

Our Approach

Preliminary Result

Conclusion

Future Work

- Experimental study is a tool to assess and compare the effectiveness of different techniques
- Experimental study needs prepared subjects
- Preparing subjects in software testing needs:
  1. A program
  2. Some known faulty versions
  3. A relatively large number of test cases

Outline
Motivation
Challenge
Motivation
Mutating the source code
Problem Statement
Possible Solution
Our Approach
Preliminary Result
Conclusion
Future Work

Software and Experiments

Motivation

Outline
Motivation
Challenge
Motivation
Mutating the source code
Problem Statement
Possible Solution
Our Approach
Preliminary Result
Conclusion
Future Work

- Problem: In software testing research there are only a few well prepared subjects
  - Finding programs with known real faults is difficult
- Not enough data: making the experiments and statistical analysis less significant
- Is there any solution?

Mutating the source code

Outline
Motivation
Challenge
Motivation
Mutating the source code
Problem Statement
Possible Solution
Our Approach
Preliminary Result
Conclusion
Future Work

- Using mutation to generate faulty versions.
- What is mutation and mutation operators:

Fault - An incorrect part of a code.

- Faulty Code: *while* $(a > b)$ *do* $\{...\}$
- Correct Code: *while* $(a < b)$ *do* $\{...\}$

Mutant - Small modification of a code by mutation operator.

- Correct Code: *while* $(a > b)$ *do* $\{...\}$
- Mutated Code: *while* $(a < b)$ *do* $\{...\}$

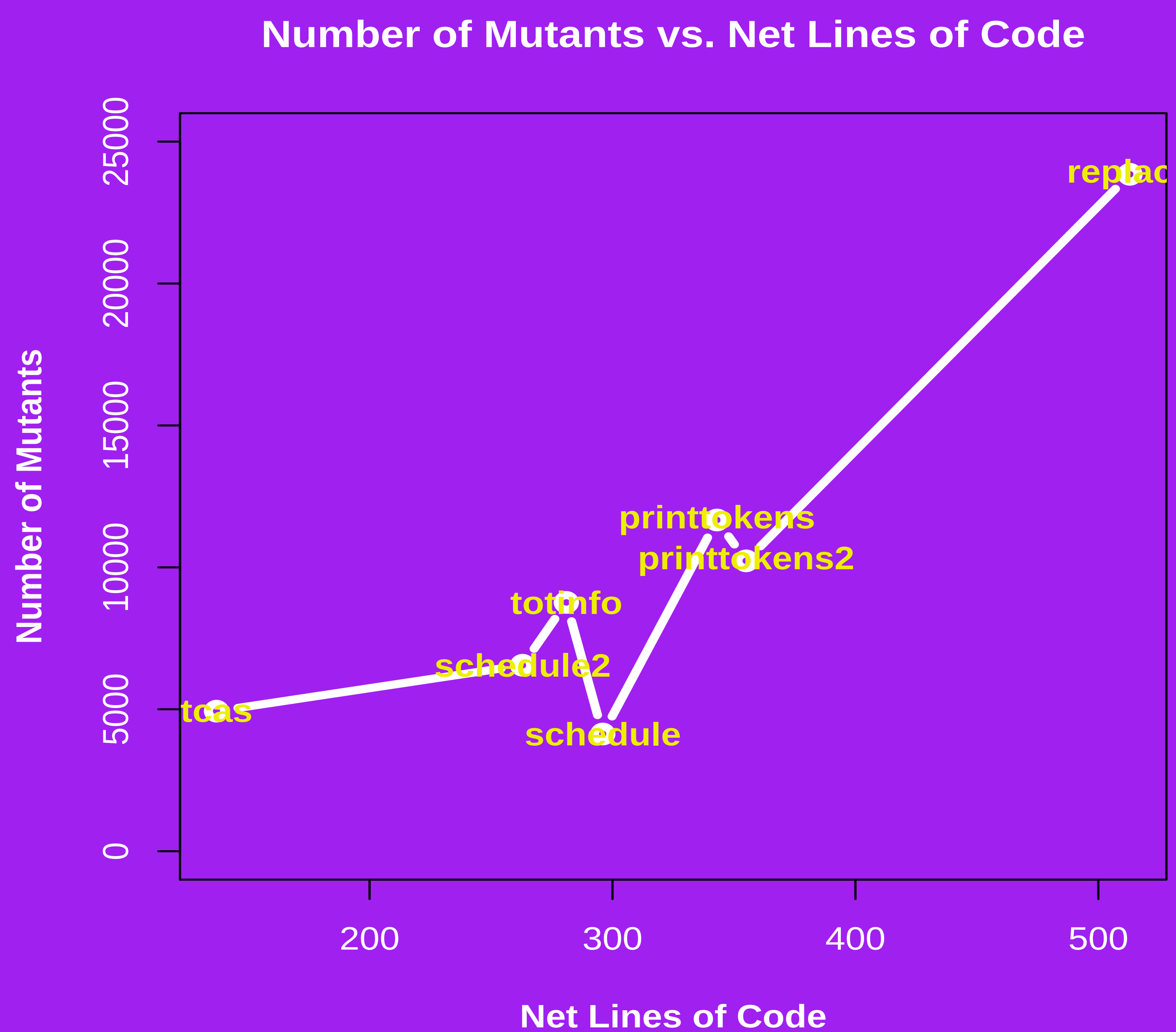
Outline
Motivation
Challenge
Problem Statement
The problem
The problem
Possible Solution
Our Approach
Preliminary Result
Conclusion
Future Work

So, what is the problem?

The Problem

Outline
Motivation
Challenge
Problem Statement
The problem
The problem
Possible Solution
Our Approach
Preliminary Result
Conclusion
Future Work

- Many mutation operators (Good or Bad?)
  - Proteum, A mutant generator for C, implements 108 operators
- Good: Simulating defects more closely
- Bad: Enormous number of mutants
  - E.g., A simple program with 137 lines of code can have 4935 mutants
- Infeasible Computation!!!


Number of Mutants vs. Net Lines of Code
Number of Mutants
0
5000
10000
15000
20000
25000
200
300
400
500
Net Lines of Code
tcas
schedule2
totinfo
schedule
printtokens
printtokens2
replac

Outline
Motivation
Challenge
Problem Statement
Possible Solution
Possible solution
Our Approach
Preliminary Result
Conclusion
Future Work

Is there any solution?

The Solution

Outline
Motivation
Challenge
Problem Statement
Possible Solution
Possible solution
Our Approach
Preliminary Result
Conclusion
Future Work

- Finding a minimal subset of mutation operators (problem?)
    - The minimal subset might not be necessarily the best
- Finding a sufficient subset of mutants called sufficient mutants (How?)
- Offutt et al. worked on selecting one of several fixed subsets of 21 mutation operators
- Wong et al. suggested defining constrained mutation so that the most critical and useful operators will be used in experiments

Outline
Motivation
Challenge
Problem Statement
Possible Solution
Our Approach
Our Approach
Metrics Correlation Analysis
Cluster Analysis
All-Subset Regression
Preliminary Result
Conclusion
Future Work

Statistical Techniques

Empirical Studies

Outline
Motivation
Challenge
Problem Statement
Possible Solution
Our Approach
Our Approach
Metrics
Correlation Analysis
Cluster Analysis
All-Subset Regression
Preliminary Result
Conclusion
Future Work

Our Approach

1. Finding sufficient subset through statistical techniques
    - Correlation Analysis
    - Cluster Analysis
    - All-Subset Regression Analysis
2. Applying multiple-linear regression on the identified enough subset

Metrics Defined

Outline
Motivation
Challenge
Problem Statement
Possible Solution
Our Approach
Our Approach
Metrics
Correlation Analysis
Cluster Analysis
All-Subset Regression
Preliminary Result
Conclusion
Future Work

- Cost of each operator μ_i:

$$C_i = \sum_j \frac{\#Mutants\ Generated\ By\ \mu_i}{\#All\ Mutants}$$

for $j \in \{Subject\ Programs\}$

- For each test suite:

1. Am_i : Detection ratio of the mutants generated by the operator μ_i
2. AM: Detection ratio of the mutants generated by all operators

Correlation Analysis: A useful technique

Outline
Motivation
Challenge
Problem Statement
Possible Solution
Our Approach
- Our Approach
- Metrics
- Correlation Analysis
- Cluster Analysis
- All-Subset Regression

Preliminary Result
Conclusion
Future Work

- Correlation between two variables (How?)
  1. Constructing correlation matrix
  2. Comparing each pair of variables
  3. Eliminating one with high cost if there is a high correlation
     - What is high correlation? $cor(X, Y) \geq 0.90$ (Standard Guilford Scalling)
- Calling the remaining variables the sufficient set
- Applying multiple-linear regression on the sufficient set

Cluster Analysis: A similarity technique

Outline
Motivation
Challenge
Problem Statement
Possible Solution
Our Approach
Our Approach
Metrics
Correlation Analysis
Cluster Analysis
All-Subset Regression
Preliminary Result
Conclusion
Future Work

- Clustering variables based on their similarities (How?)
  1. Constructing proximity matrix
     - Computing the nearest neighbor
  2. Generating a dendrogram
  3. Eliminating one with high cost if there is a high correlation
     - What is high correlation? $cor(X, Y) \geq 0.90$
- Calling the remaining variables the sufficient set
- Applying multiple-linear regression on the sufficient set

All-Subset Regression: A search

Outline

Motivation

Challenge

Problem Statement

Possible Solution

Our Approach

- Our Approach
- Metrics
- Correlation Analysis
- Cluster Analysis
- All-Subset Regression

Preliminary Result

Conclusion

Future Work

- Select a small subset of a large set of variables that model a response variable (How?)
  1. Constructing all regression models exhaustively
     - Number of variables involved in each model chosen by the analyst
  2. The result is already a multiple-linear regression
- Calling the variables involved in the chosen model the sufficient set

Outline
Motivation
Challenge
Problem Statement
Possible Solution
Our Approach
Preliminary Result
Preliminary Result
Preliminary Result
Preliminary Result
Preliminary Result
Conclusion
Future Work

Preliminary Result

Preliminary Result

Outline
Motivation
Challenge
Problem Statement
Possible Solution
Our Approach
Preliminary Result
Preliminary Result
Preliminary Result
Preliminary Result
Preliminary Result
Conclusion
Future Work

- Applied the techniques to tcas.c program so far
- 300 test suites for each group
- Applied only to random-based test suites

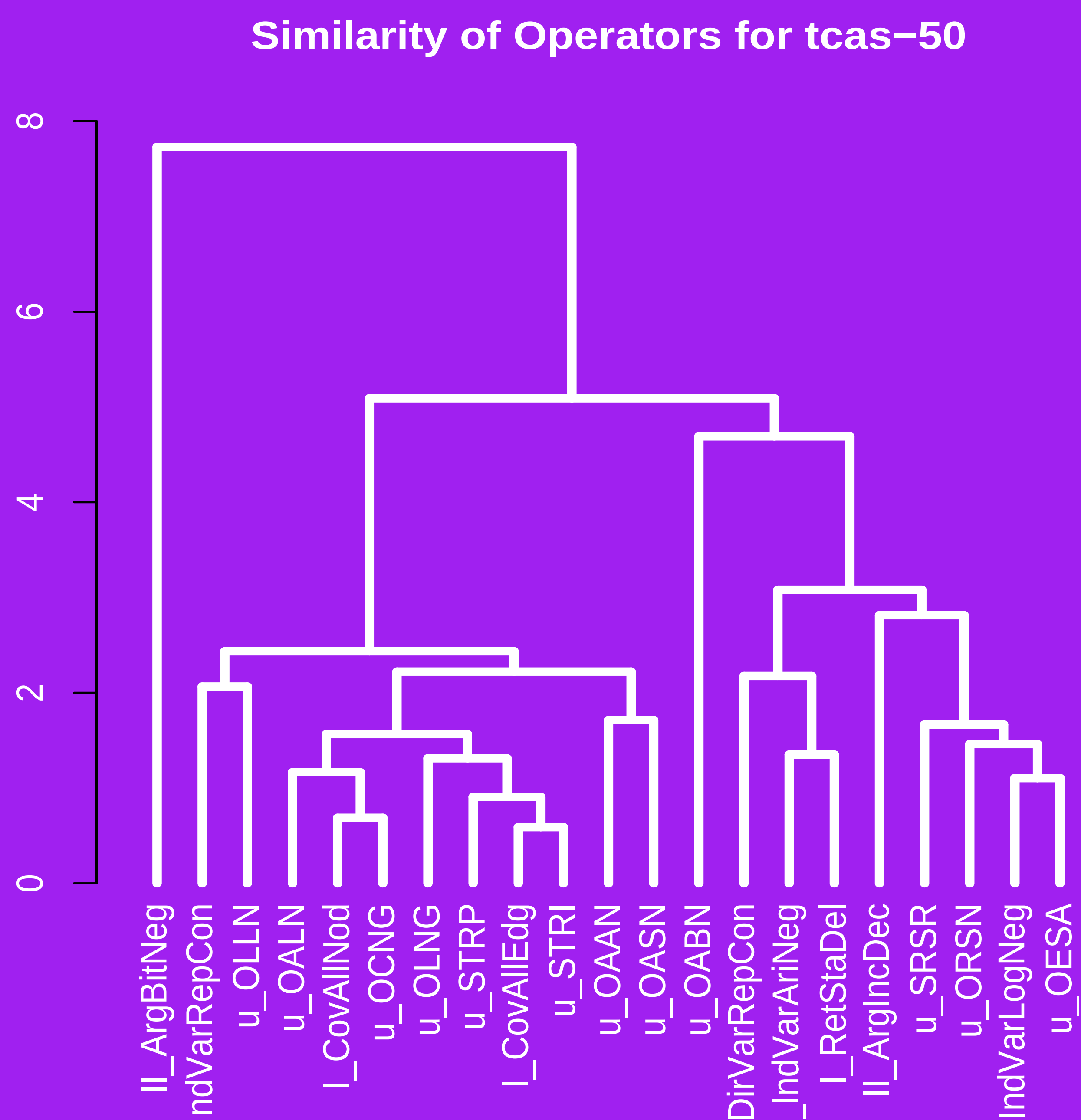
Preliminary Result

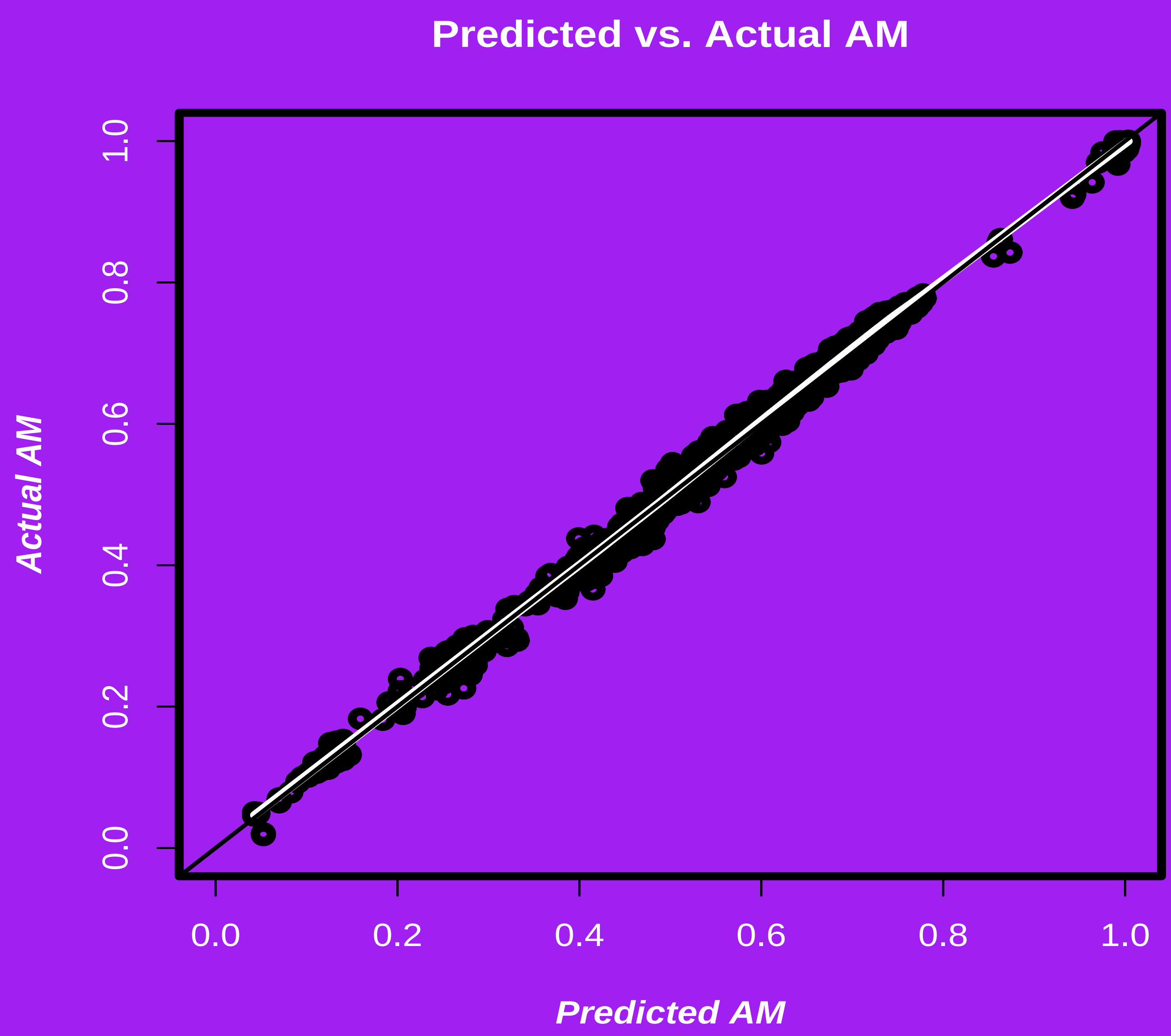
Outline
Motivation
Challenge
Problem Statement
Possible Solution
Our Approach
Preliminary Result
Preliminary Result
Preliminary Result
Preliminary Result
Preliminary Result
Conclusion
Future Work

core-SUB	Mutants
I-DirVarAriNeg	44
I-DirVarRepReq	220
I-IndVarLogNeg	19
I-IndVarRepReq	91
u-OABN	3
u-OLSN	34
u-VDTR	111
u-VGSR	794
u-VTWD	74
II-ArgLogNeg	3
Total	1393

core-EBC	Mutants
I-IndVarAriNeg	19
I-IndVarLogNeg	19
II-ArgRepReq	5
u-OALN	2
u-OASN	2
u-OCNG	8
u-OLLN	17
Total	72

core-CA	Mutants
I-CovAllEdg	12
I-DirVarRepCon	38
I-IndVarAriNeg	19
I-RetStaDel	17
II-ArgIncDec	54
u-OABN	3
u-OALN	2
u-OASN	2
u-OCNG	8
u-OLNG	51
u-ORSN	30
u-SRSR	60
u-STRP	70
Total	366


Similarity of Operators for tcas−50
0
2
4
6
8
II_ArgBitNeg
ndVarRepCon
u_OLLN
u_OALN
I_CovAllNod
u_OCNG
u_OLNG
u_STRP
I_CovAllEdg
u_STRI
u_OAAN
u_OASN
u_OABN
DirVarRepCon
IndVarAriNeg
I_RetStaDel
II_ArgIncDec
u_SRSR
u_ORSN
IndVarLogNeg
u_OESA



Predicted vs. Actual AM
Actual AM
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
Predicted AM

Outline
Motivation
Challenge
Problem Statement
Possible Solution
Our Approach
Preliminary Result
Conclusion
Conclusion
Future Work

Conclusion

Conclusion

Outline
Motivation
Challenge
Problem Statement
Possible Solution
Our Approach
Preliminary Result
Conclusion
Conclusion
Future Work

- Addressing the expensive cost of mutation by three statistical analysis
  - Correlation Analysis
  - Cluster Analysis
  - All-Subset Regression Analysis
- Which one is better?
  - In terms of efficiency: Cluster Analysis
  - In terms of accuracy: All-Subset Regression Analysis

Outline
Motivation
Challenge
Problem Statement
Possible Solution
Our Approach
Preliminary Result
Conclusion
Future Work
Future Work

Future Work

Future Work

Outline
Motivation
Challenge
Problem Statement
Possible Solution
Our Approach
Preliminary Result
Conclusion
Future Work
Future Work

- Applying techniques to all Siemens subject programs
- Verifying the result with other subject programs
- Applying the techniques to different programming languages using different mutant generators