

Introduction

Introduction

Problem

Pseudo-tested
Methods

Function-
Fiasco

Evaluation
Strategy

Conclusion

Automatic Detection of Pseudo-Tested Methods in a Test Suite Using Fault Injection

Nicholas Tocci

November 19, 2018

Problem

Introduction

Problem

Coverage

Pseudo-tested
Methods

Function-
Fiasco

Evaluation
Strategy

Conclusion

How can we know if our test suites
are adequate?



Coverage

Introduction

Problem

Coverage

Calculation

Coverage vs

Adequate

Coverage

Pseudo-tested

Methods

Function-

Fiasco

Evaluation

Strategy

Conclusion

Coverage



Def: % of a system that has been tested.

Calculation

Introduction

Problem

Coverage

Calculation

Coverage vs

Adequate

Coverage

Pseudo-tested
Methods

Function-
Fiasco

Evaluation
Strategy

Conclusion

$$\text{Coverage} = \frac{\text{Number of Tested Methods}}{\text{Total Number of Methods}}$$

High Coverage

Introduction

Problem

Coverage

Calculation

Coverage vs
Adequate
Coverage

Pseudo-tested
Methods

Function-
Fiasco

Evaluation
Strategy

Conclusion



Pseudo-tested Methods

Introduction

Problem

**Pseudo-tested
Methods**

Defintion
Detection

Function-
Fiasco

Evaluation
Strategy

Conclusion

Pseudo-tested Methods

What is a Pseudo-tested Method?

PASSED

Def: It will never fail.

How Can We Detect Pseudo-tested Methods

It is harder than you think!

Example of a Pseudo-tested method

Introduction

Problem

Pseudo-tested

Methods

Definition

Detection

Function-

Fiasco

Evaluation

Strategy

Conclusion

```
numbers.py:
def numberOrder(n):
    numbersSorted = sorted(n)
    return numbersSorted

test_numbers.py:
def test_numbers_ordered():
    numbers = {2,4,3,1}
    sortedNumbers = {1,2,3,4}
    orderedNumbers = numberOrder(numbers)
    assert numbers == sortedNumbers
```

What is Function-Fiasco

Introduction
Problem
Pseudo-tested
Methods
Function-
Fiasco
What is
Function-Fiasco
Flow
Evaluation
Strategy
Conclusion

A Pseudo-tested method detection tool



Flow to system

Introduction

Problem

Pseudo-tested
Methods

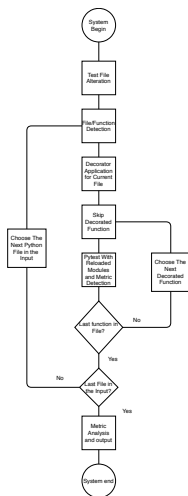
Function-
Fiasco

What is
Function-Fiasco

Flow

Evaluation
Strategy

Conclusion



Beginning

Introduction

Problem

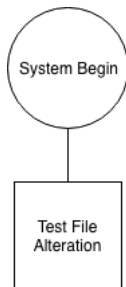
Pseudo-tested
Methods

Function-
Fiasco

What is
Function-Fiasco
Flow

Evaluation
Strategy

Conclusion



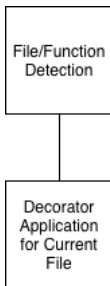
Reloader Function

- Introduction
- Problem
- Pseudo-tested
Methods
- Function-
Fiasco
 - What is
Function-Fiasco
- Flow
- Evaluation
Strategy
- Conclusion

```
def reloader():  
    for modules in listToReload:  
        importlib.reload(modules)
```

Next Steps

Introduction
Problem
Pseudo-tested
Methods
Function-
Fiasco
What is
Function-Fiasco
Flow
Evaluation
Strategy
Conclusion



Decorator Function

Introduction

Problem

Pseudo-tested
Methods

Function-
Fiasco

What is
Function-Fiasco
Flow

Evaluation
Strategy

Conclusion

```
def skipper(func):  
    if(func.__name__ in functionsComplete):  
        def doFunc(*args , **kwargs):  
            return 5  
        return doFunc  
  
    else:  
        def doFunc(*args , **kwargs):  
            var = func(*args , **kwargs)  
            return var  
        return doFunc
```

Pytesting

Introduction

Problem

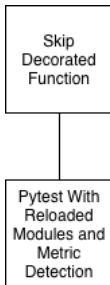
Pseudo-tested
Methods

Function-
Fiasco

What is
Function-Fiasco
Flow

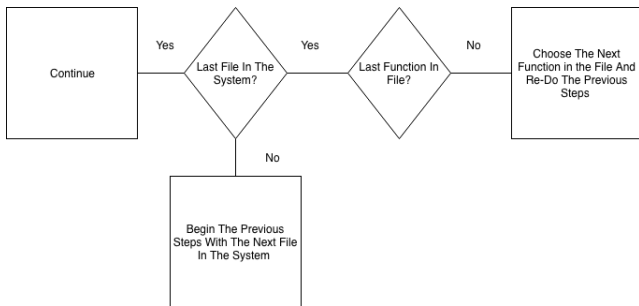
Evaluation
Strategy

Conclusion



File and Function Checks

Introduction
Problem
Pseudo-tested
Methods
Function-
Fiasco
What is
Function-Fiasco
Flow
Evaluation
Strategy
Conclusion



System End

Introduction

Problem

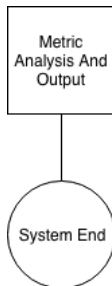
Pseudo-tested
Methods

Function-
Fiasco

What is
Function-Fiasco
Flow

Evaluation
Strategy

Conclusion



Coverage Calculation

Introduction

Problem

Pseudo-tested
Methods

Function-
Fiasco

Evaluation
Strategy

Coverage
Calculation

Truly-Tested-
Method
Calculation

Metrics
Produced

Conclusion

$$\text{Coverage} = \frac{\text{Number of Tested Methods}}{\text{Total Number of Methods}}$$

Coverage Example

Introduction

Problem

Pseudo-tested
Methods

Function-
Fiasco

Evaluation
Strategy

Coverage
Calculation

Truly- Tested-
Method
Calculation

Metrics
Produced

Conclusion

| NUMM | NUMTM | Coverage |
|------|-------|----------|
| 40 | 25 | 62.5% |

Truly-Tested-Method Calculation

Introduction

Problem

Pseudo-tested
Methods

Function-
Fiasco

Evaluation
Strategy

Coverage
Calculation

Truly-Tested-
Method
Calculation

Metrics
Produced

Conclusion

- **Number of Truly-Tested-Methods = NUMTTM**
- **Number of Tested Methods = NUMTM**
- **Number of Pseudo-tested Methods = NUMPTM**

$$NUMTTM = NUMTM - NUMPTM$$

Truly-Tested-Method Example

Introduction

Problem

Pseudo-tested
Methods

Function-
Fiasco

Evaluation
Strategy

Coverage
Calculation

Truly-Tested-
Method
Calculation

Metrics
Produced

Conclusion

| NUMTM | NUMPTM | NUMTTM |
|-------|--------|--------|
| 25 | 3 | 22 |

Adequate-Coverage Calculation

Introduction
Problem
Pseudo-tested
Methods
Function-
Fiasco
Evaluation
Strategy
Coverage
Calculation
Truly-Tested-
Method
Calculation
Metrics
Produced
Conclusion

$$AC = \frac{\textit{NumberofTrulyTestedMethods}}{\textit{TotalNumberofMethods}}$$

Output

Introduction

Problem

Pseudo-tested
Methods

Function-
Fiasco

Evaluation
Strategy

Coverage
Calculation

Truly- Tested-
Method
Calculation

Metrics
Produced

Conclusion

| NUMM | NUMTM | Coverage | NUMPTM | NUMTTM | AC |
|------|-------|----------|--------|--------|-----|
| 40 | 25 | 62.5% | 3 | 22 | 55% |

What is different?

Introduction

Problem

Pseudo-tested
Methods

Function-
Fiasco

Evaluation
Strategy

Conclusion

**What is
different?**

Why

Impact

Feasibility

Future Research

Demo

FFiasco VS Mutation VS Fuzz VS Fault Injection VS Chaos

Why is a tool like this necessary?

Introduction

Problem

Pseudo-tested
Methods

Function-
Fiasco

Evaluation
Strategy

Conclusion

What is
different?

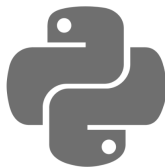
Why

Impact

Feasibility

Future Research

Demo



- **Nothing like this exists for Python**
- **Coverage does not tell enough**

What is the impact of this research?

Introduction

Problem

Pseudo-tested
Methods

Function-
Fiasco

Evaluation
Strategy

Conclusion

What is
different?

Why

Impact

Feasibility

Future Research

Demo



Coverage with
fault detection



Better
Understanding of
Pseudo-tested
Methods



Automatic
Detection Tool

Feasibility

Introduction

Problem

Pseudo-tested
Methods

Function-
Fiasco

Evaluation
Strategy

Conclusion

What is
different?

Why

Impact

Feasibility

Future Research

Demo

Github and GatorGrader

Future Research

Introduction

Problem

Pseudo-tested
Methods

Function-
Fiasco

Evaluation
Strategy

Conclusion

What is
different?

Why
Impact
Feasibility

Future Research

Demo

- Return Inferencing
- Hypothesis
- Pytest Plugins



Demo

Introduction

Problem

Pseudo-tested
Methods

Function-
Fiasco

Evaluation
Strategy

Conclusion

What is
different?

Why

Impact

Feasibility

Future Research

Demo

DEMO