# Fine-Tuning Spectrum Based Fault Localisation with Sequence Mining

**Gulsher Laghari**, Alessandro Murgia and Serge Demeyer

BENEVOL2016 - December 8, 2016

Ansymo
Antwerp Systems & Software Modelling
University of Antwerp

Universiteit Antwerpen

# Overview


CHANGES
CHANGES
CHANGES
CHANGES


class
class
class Token{
char
char

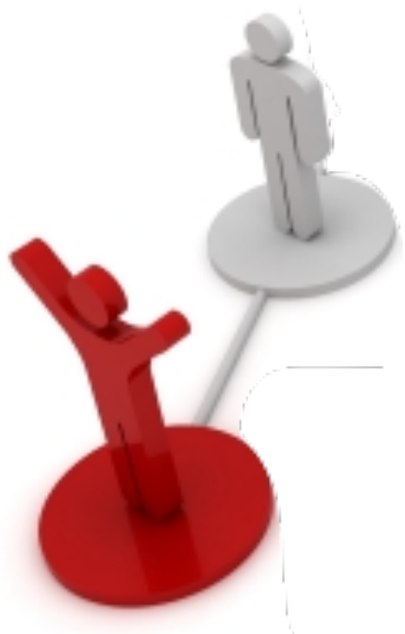
# Fault Localisation

Locate Understand Fix

Fault Localisation an important step in debugging process

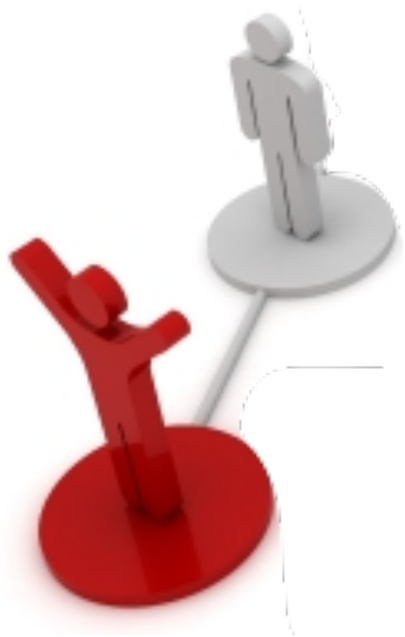
# Fault Localisation

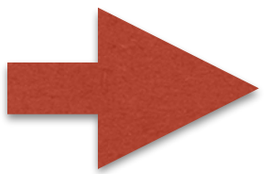
## Test to code mapping



# Fault Localisation

## Test to code mapping





# Fault Localisation

## Test to code mapping

# Fault Localisation

```
public void testGetMaxMiddleIndex() {
        TimePeriodValues s = new TimePeriodValues("Test");
        assertEquals(-1, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(100L, 200L), 1.0);
        assertEquals(0, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(300L, 400L), 2.0);
        assertEquals(1, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(0L, 50L), 3.0);
        assertEquals(1, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(150L, 200L), 4.0);
        assertEquals(1, s.getMaxMiddleIndex());
    }
```

* Failing test in Apache Commons Math

# Fault Localisation

```
public void testGetMaxMiddleIndex() {
        TimePeriodValues s = new TimePeriodValues("Test");
        assertEquals(-1, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(100L, 200L), 1.0);
        assertEquals(0, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(300L, 400L), 2.0);
        assertEquals(1, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(0L, 50L), 3.0);
        assertEquals(1, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(150L, 200L), 4.0);
        assertEquals(1, s.getMaxMiddleIndex());
    }
```

* Failing test in Apache Commons Math

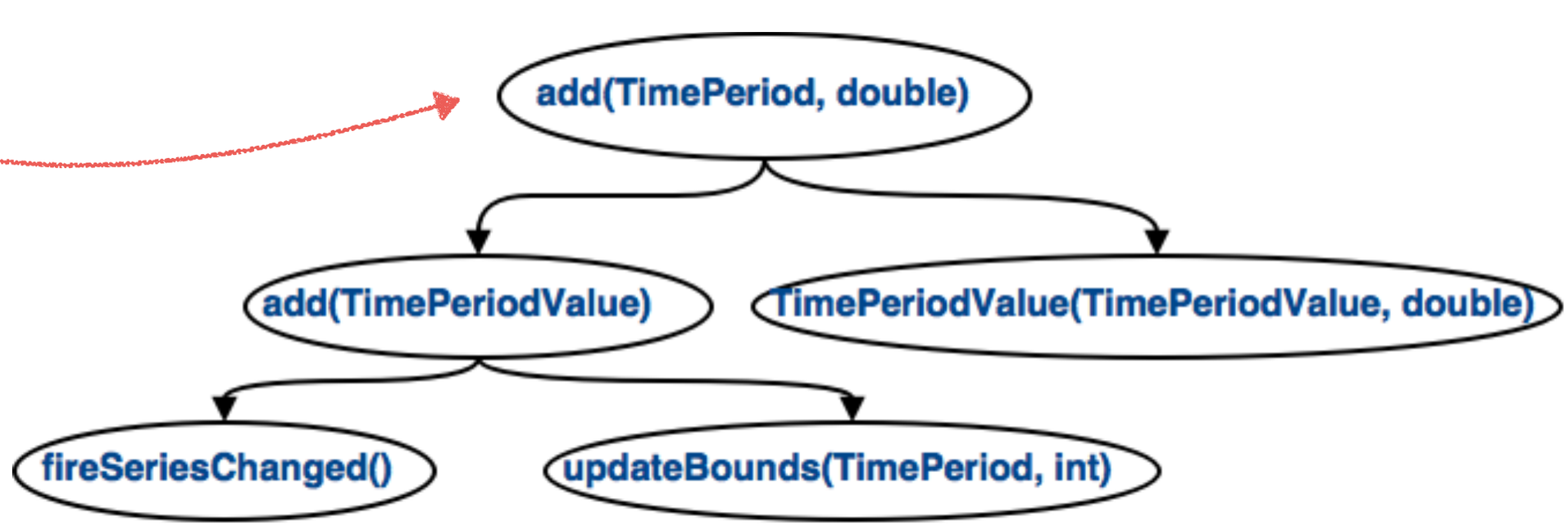
# Fault Localisation

```
public void testGetMaxMiddleIndex() {
        TimePeriodValues s = new TimePeriodValues("Test");
        assertEquals(-1, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(100L, 200L), 1.0);
        assertEquals(0, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(300L, 400L), 2.0);
        assertEquals(1, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(0L, 50L), 3.0);
        assertEquals(1, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(150L, 200L), 4.0);
        assertEquals(1, s.getMaxMiddleIndex());
    }
```

* Failing test in Apache Commons Math

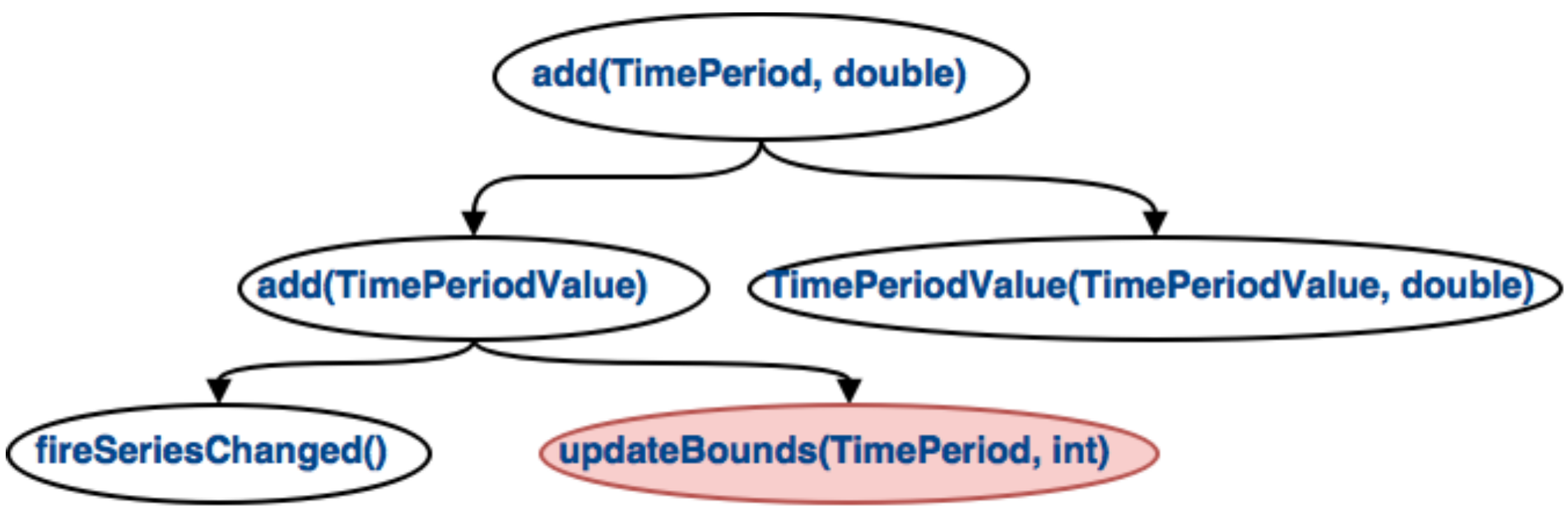
# Fault Localisation

```
public void testGetMaxMiddleIndex() {
        TimePeriodValues s = new TimePeriodValues("Test");
        assertEquals(-1, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(100L, 200L), 1.0);
        assertEquals(0, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(300L, 400L), 2.0);
        assertEquals(1, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(0L, 50L), 3.0);
        assertEquals(1, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(150L, 200L), 4.0);
        assertEquals(1, s.getMaxMiddleIndex());
    }
```



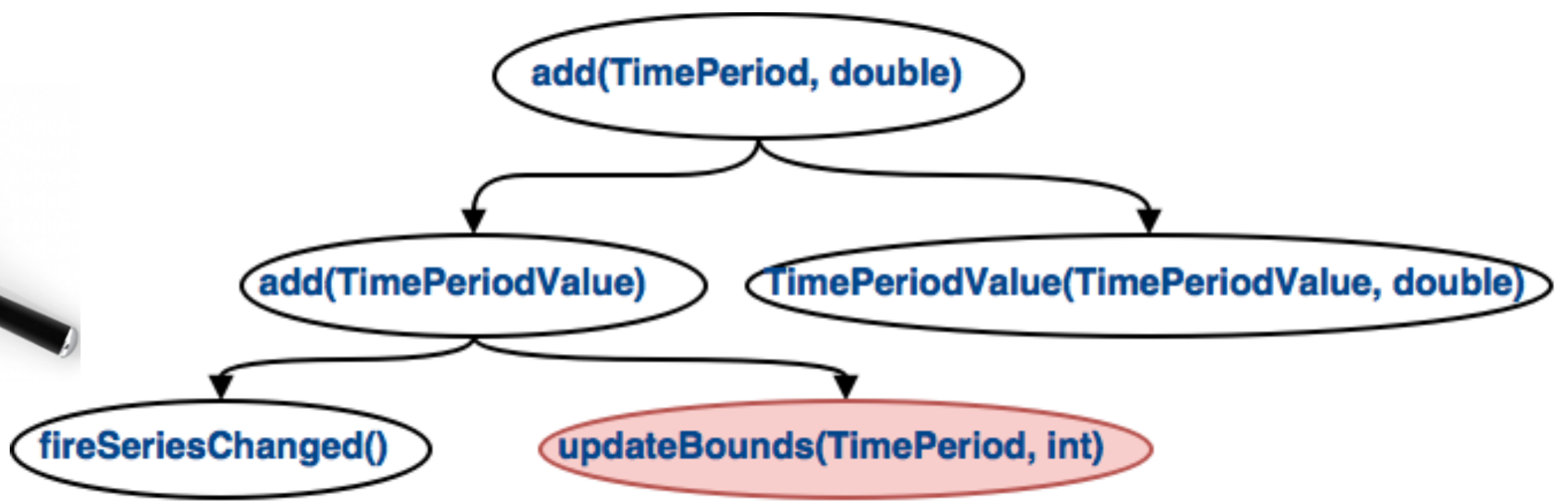
# Fault Localisation

```
public void testGetMaxMiddleIndex() {
        TimePeriodValues s = new TimePeriodValues("Test");
        assertEquals(-1, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(100L, 200L), 1.0);
        assertEquals(0, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(300L, 400L), 2.0);
        assertEquals(1, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(0L, 50L), 3.0);
        assertEquals(1, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(150L, 200L), 4.0);
        assertEquals(1, s.getMaxMiddleIndex());
    }
```



# Fault Localisation

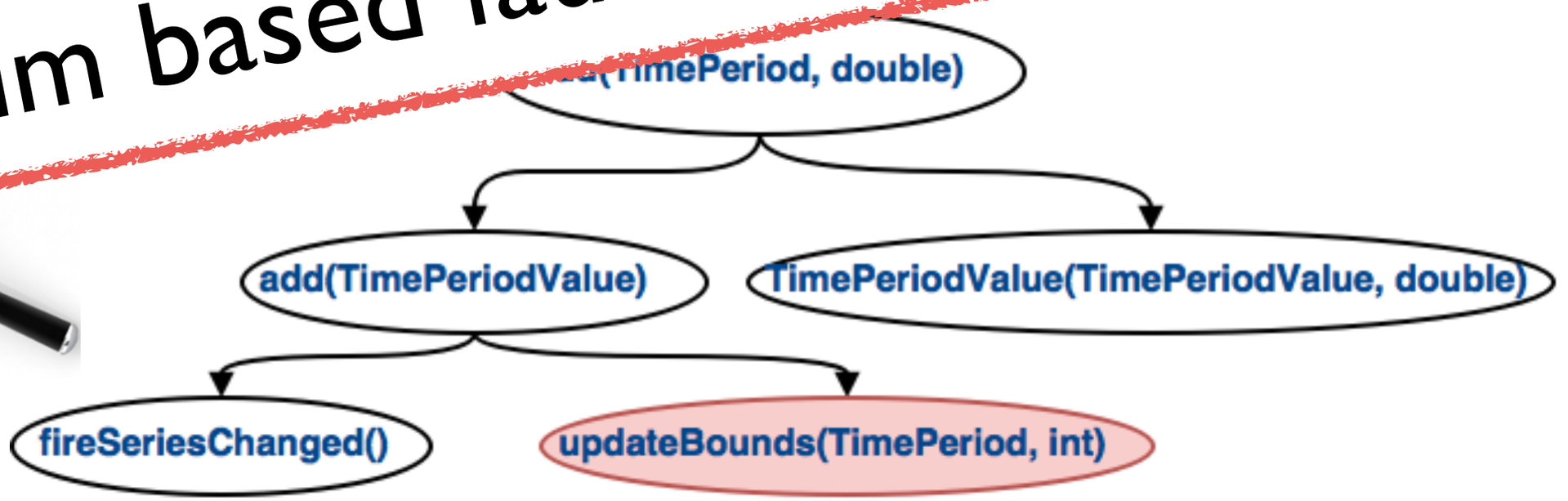
```
public void testGetMaxMiddleIndex() {
        TimePeriodValues s = new TimePeriodValues("Test");
        assertEquals(-1, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(100L, 200L), 1.0);
        assertEquals(0, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(300L, 400L), 2.0);
        assertEquals(1, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(0L, 50L), 3.0);
        assertEquals(1, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(150L, 200L), 4.0);
        assertEquals(1, s.getMaxMiddleIndex());
    }
```



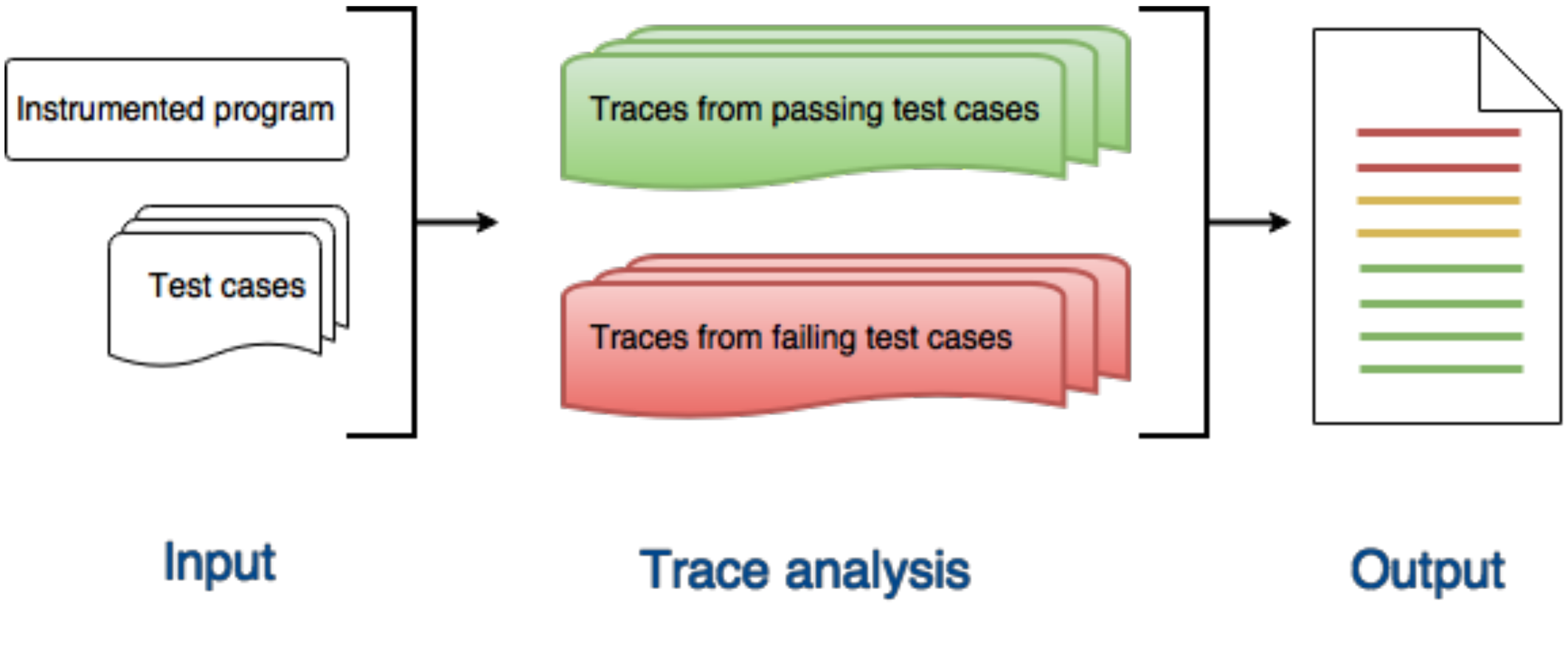
# Fault Localisation

```
public void testGetMaxMiddleIndex() {
        TimePeriodValues s = new TimePeriodValues("Test");
        assertEquals(-1, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(100L, 200L), 1.0);
        assertEquals(0, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(300L, 400L), 2.0);
        assertEquals(1, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(0L, 50L), 3.0);
        assertEquals(1, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(150L, 200L), 4.0);
        assertEquals(1, s.getMaxMiddleIndex());
    }
```

Spectrum based fault localisation

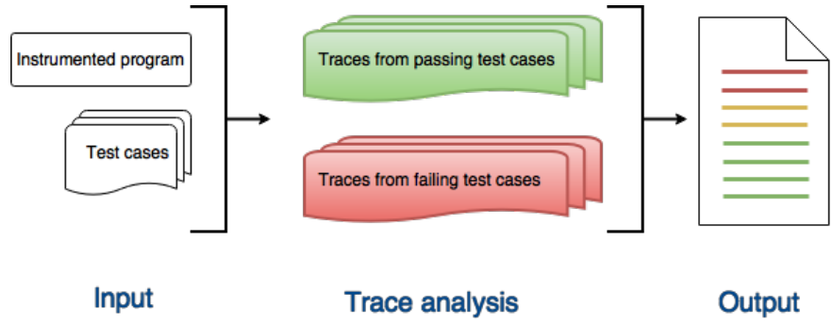


# Spectrum Based Fault Localisation



# Spectrum Based Fault Localisation

## Granularity



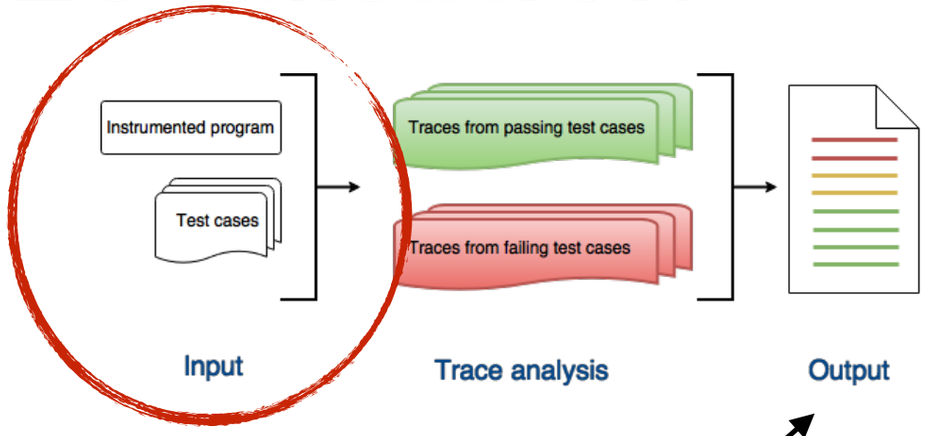

Method

Block

Statement

# Spectrum Based Fault Localisation

## Granularity



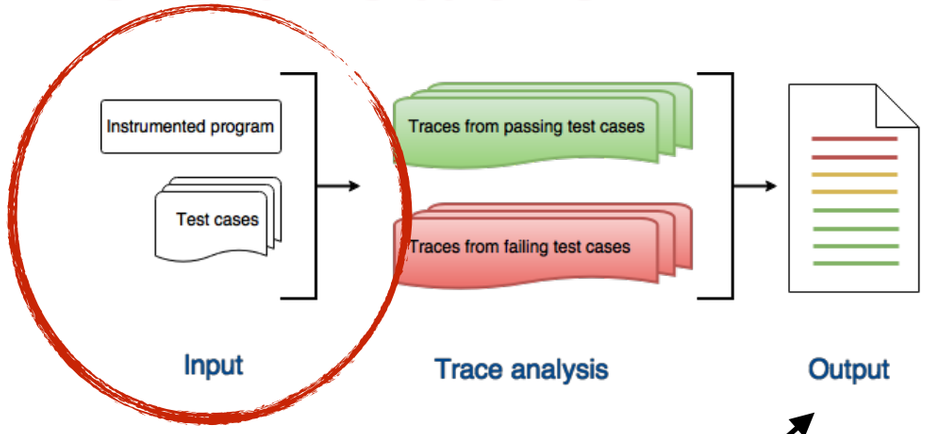

Method

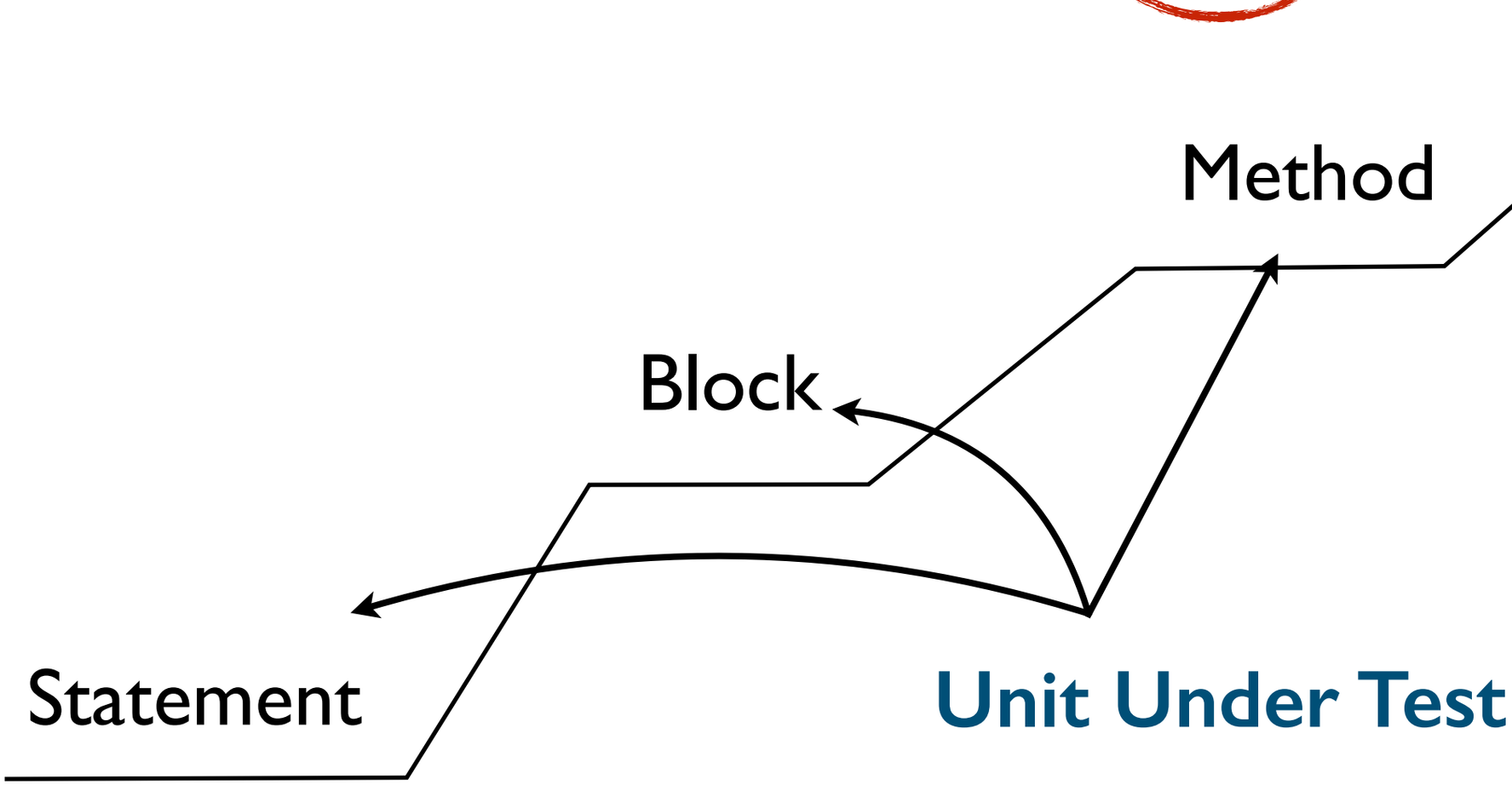
Block

Statement

# Spectrum Based Fault Localisation

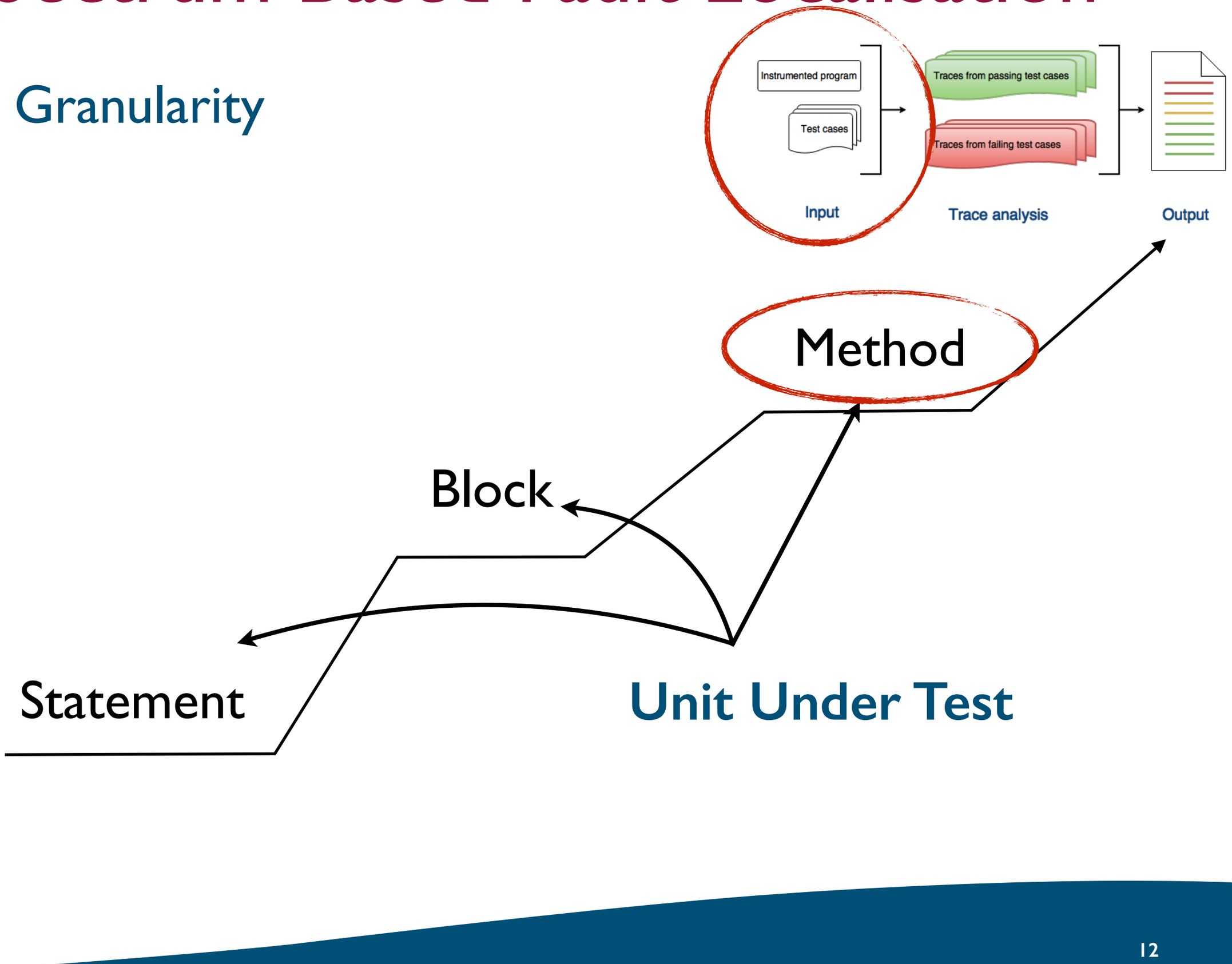
## Granularity



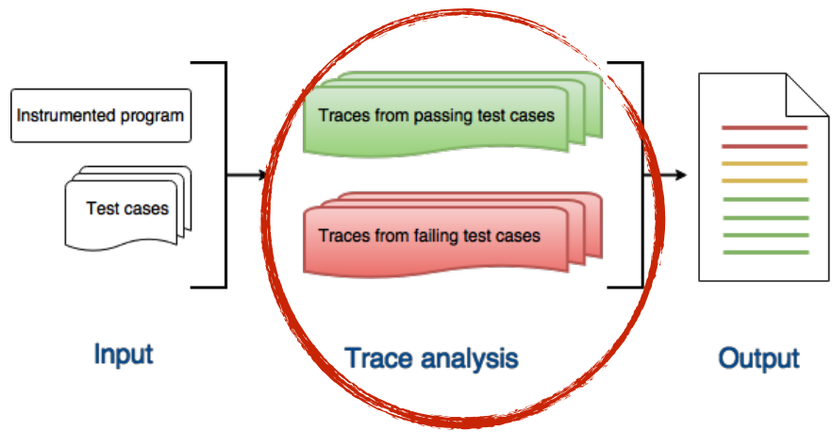
# Spectrum Based Fault Localisation

## Granularity

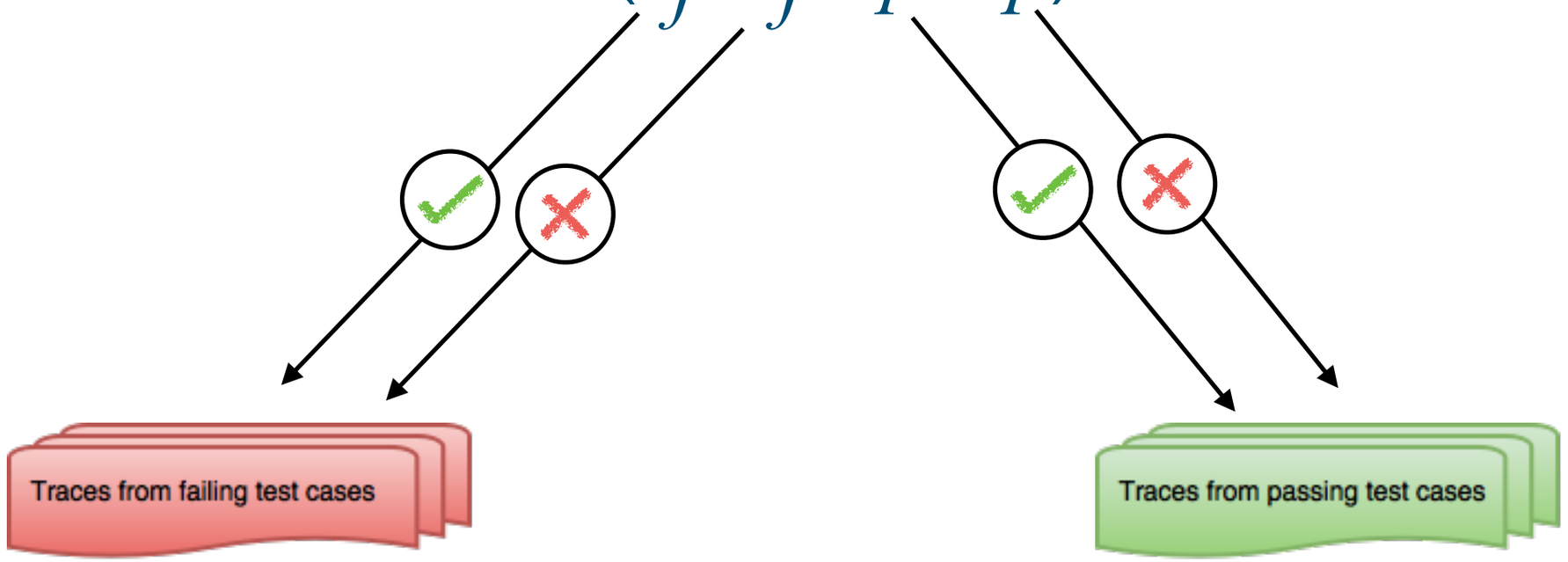


# Spectrum Based Fault Localisation

## Hit Spectrum



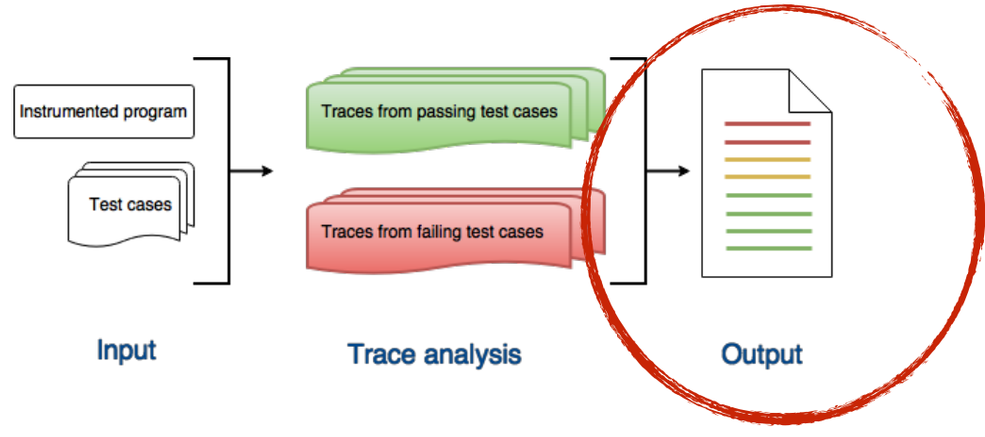

UUT = $(e_f, n_f, e_p, n_p)$




- ✓ Number of traces that contain UUT
- ✗ Number of traces that do not contain UUT

# Spectrum Based Fault Localisation

## Fault Locator




Fault Locator

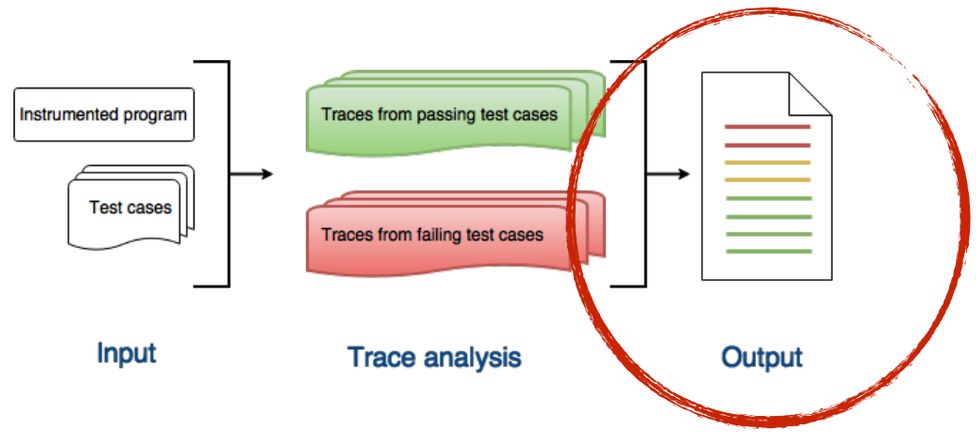
UUT = $(e_f, n_f, e_p, n_p)$

UUT = Suspiciousness

$$suspiciousness = \frac{e_f}{\sqrt{(e_f + n_f)(e_f + e_p)}}$$

# Spectrum Based Fault Localisation

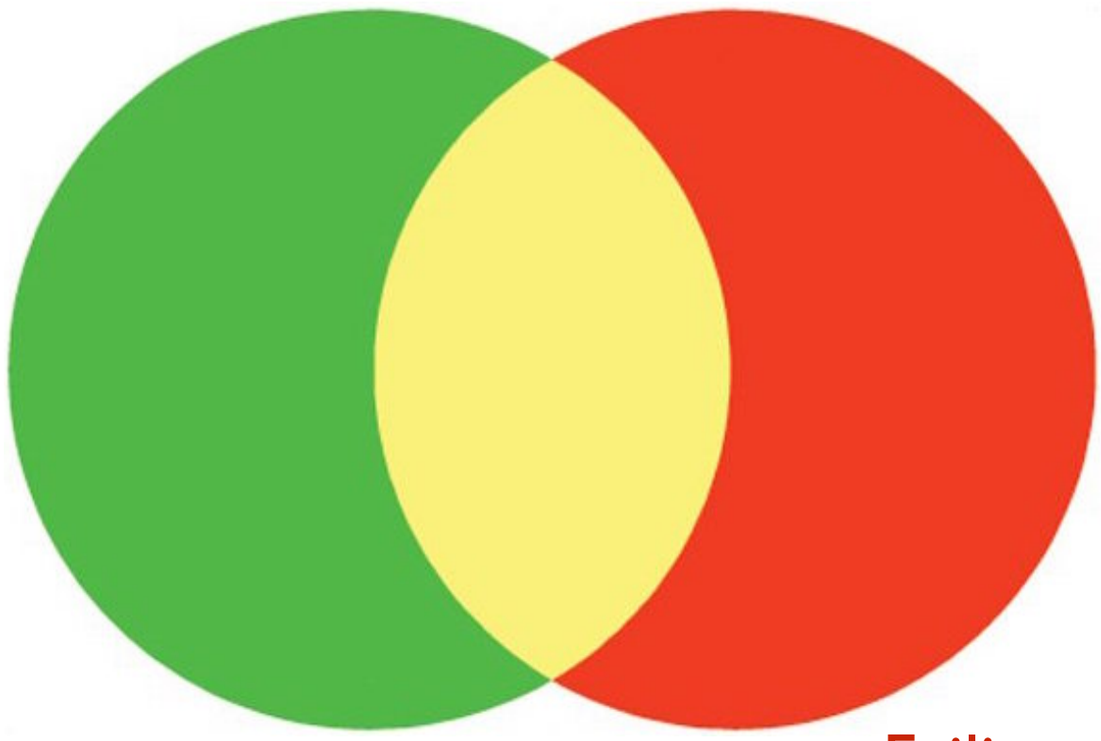
## Fault Locator




Suspiciousness = [0,1]

Not suspicious → Highly suspicious

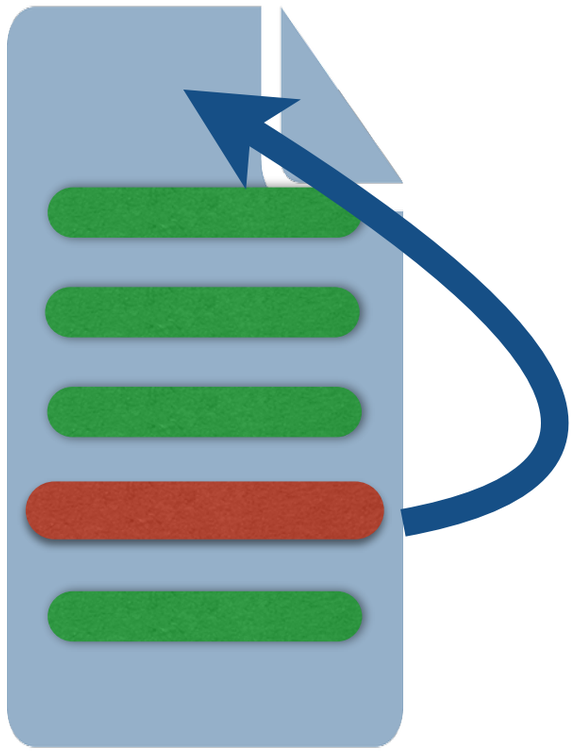
0 — 1



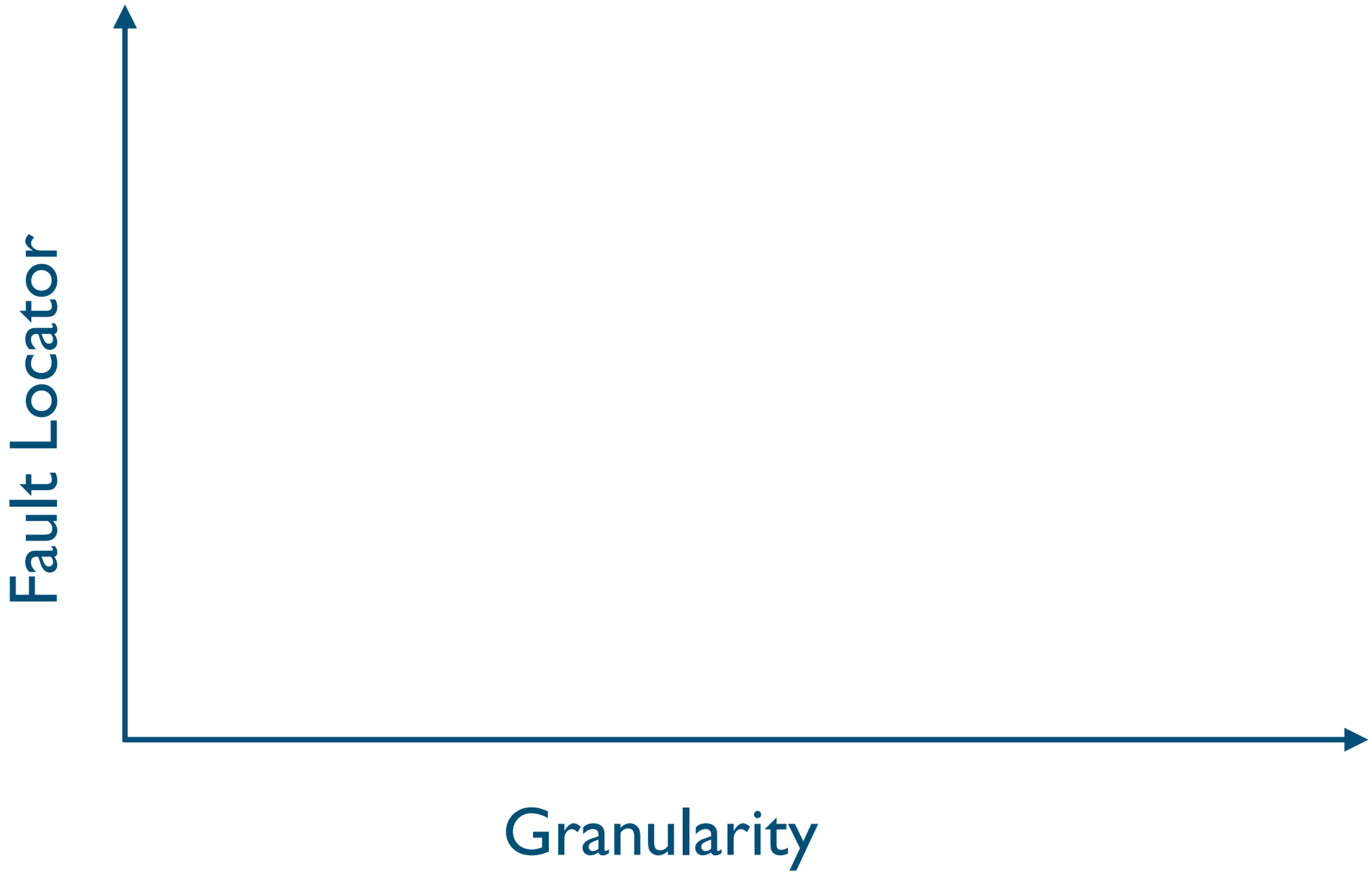
Passing test cases

Failing test cases

# Spectrum Based Fault Localisation

# Spectrum Based Fault Localisation



# Spectrum Based Fault Localisation

# Raw Spectrum Analysis

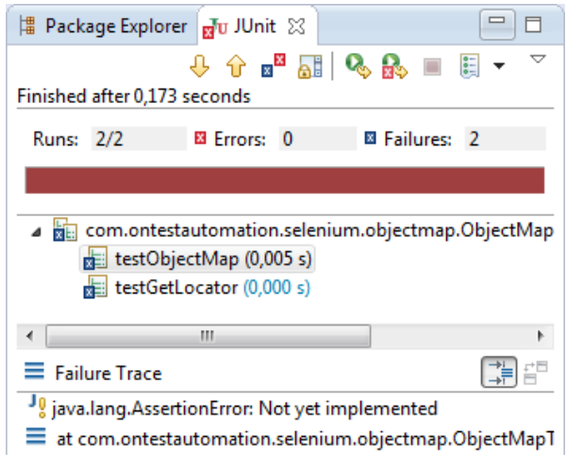
## Missing

```
method1(){
        methodA()
        methodB()
        if(condition){
          return
        }
        methodC()
}
```

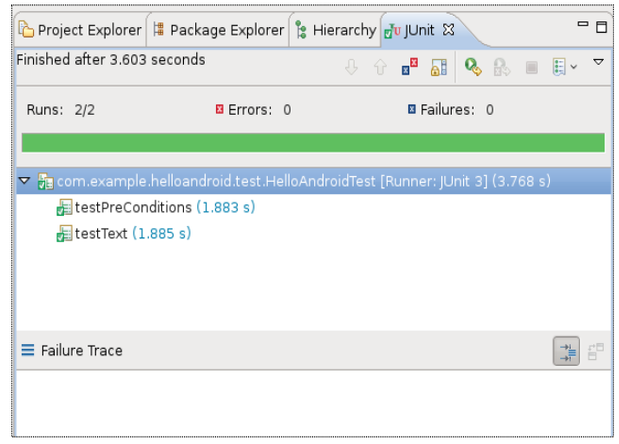
Fault

# Raw Spectrum Analysis

## Missing



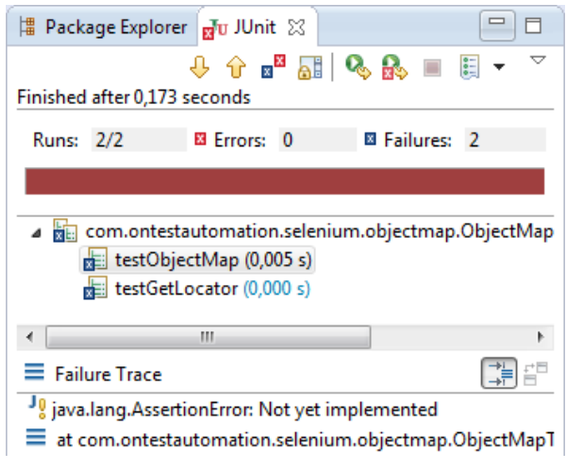

```
method1(){
  methodA()
  methodB()
  if(condition){
     return
  }
  methodC()
}
```



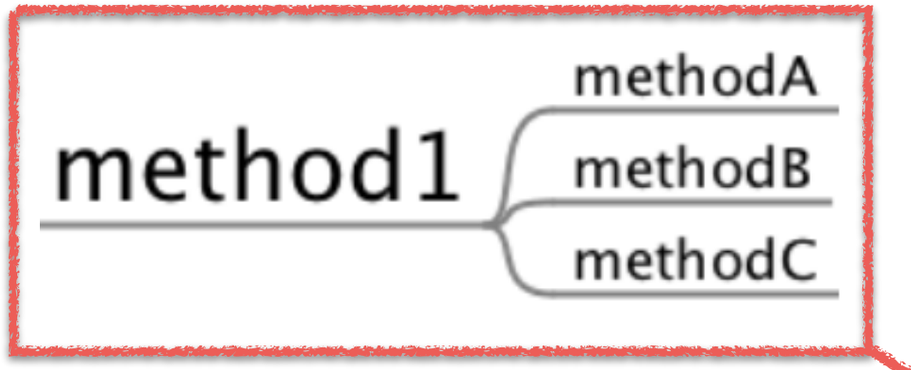

```
method1(){
  methodA()
  methodB()
  if(condition){
     return
  }
  methodC()
}
```

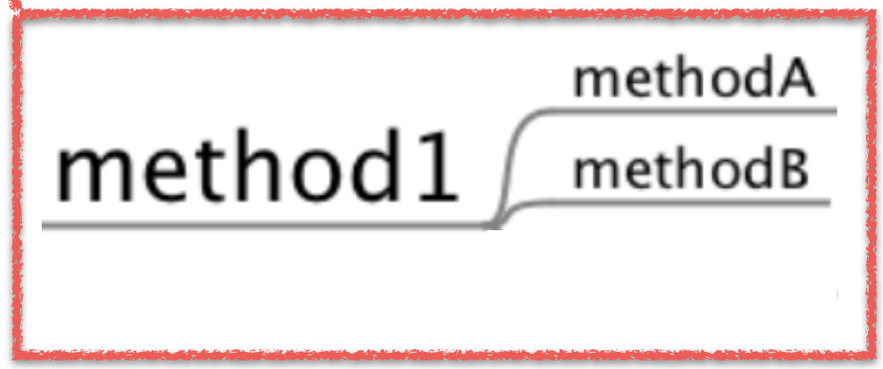
# Raw Spectrum Analysis

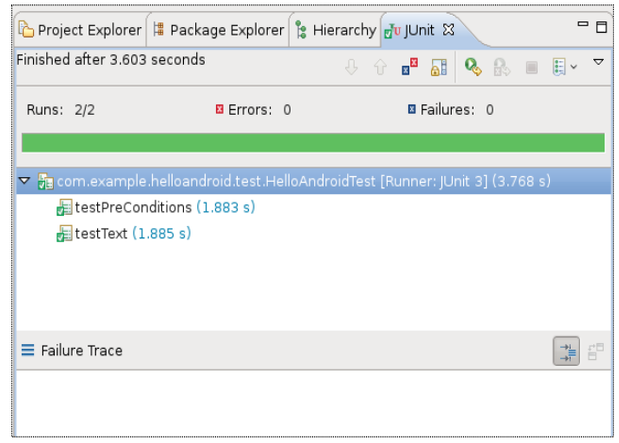
## Missing




```
method1(){
  methodA()
  methodB()
  if(condition){
     return
  }
  methodC()
}
```



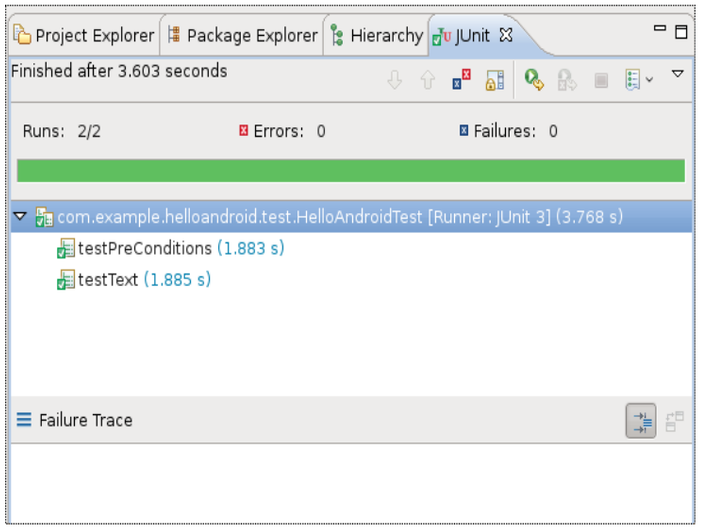



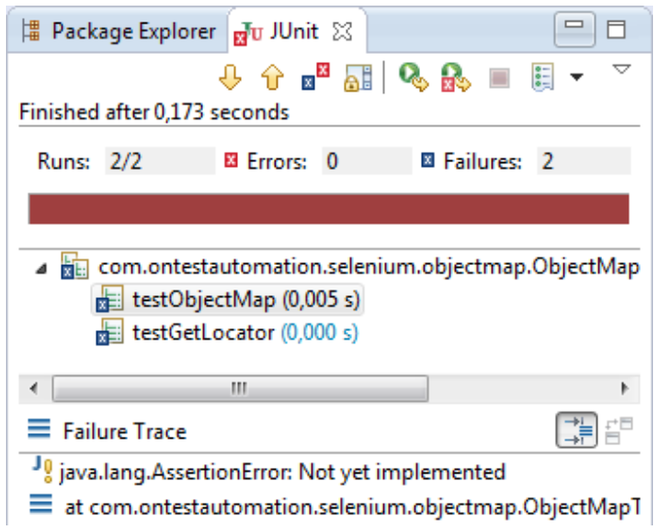



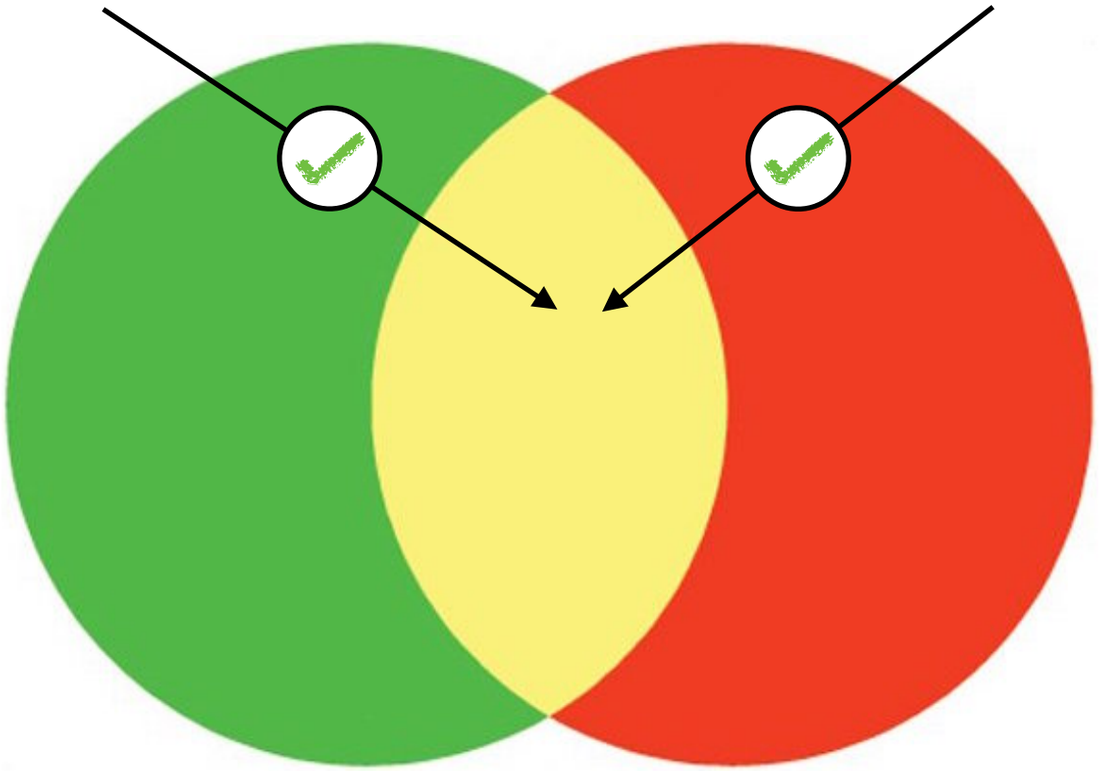

```
method1(){
  methodA()
  methodB()
  if(condition){
     return
  }
  methodC()
}
```

# Raw Spectrum Analysis

## Missing



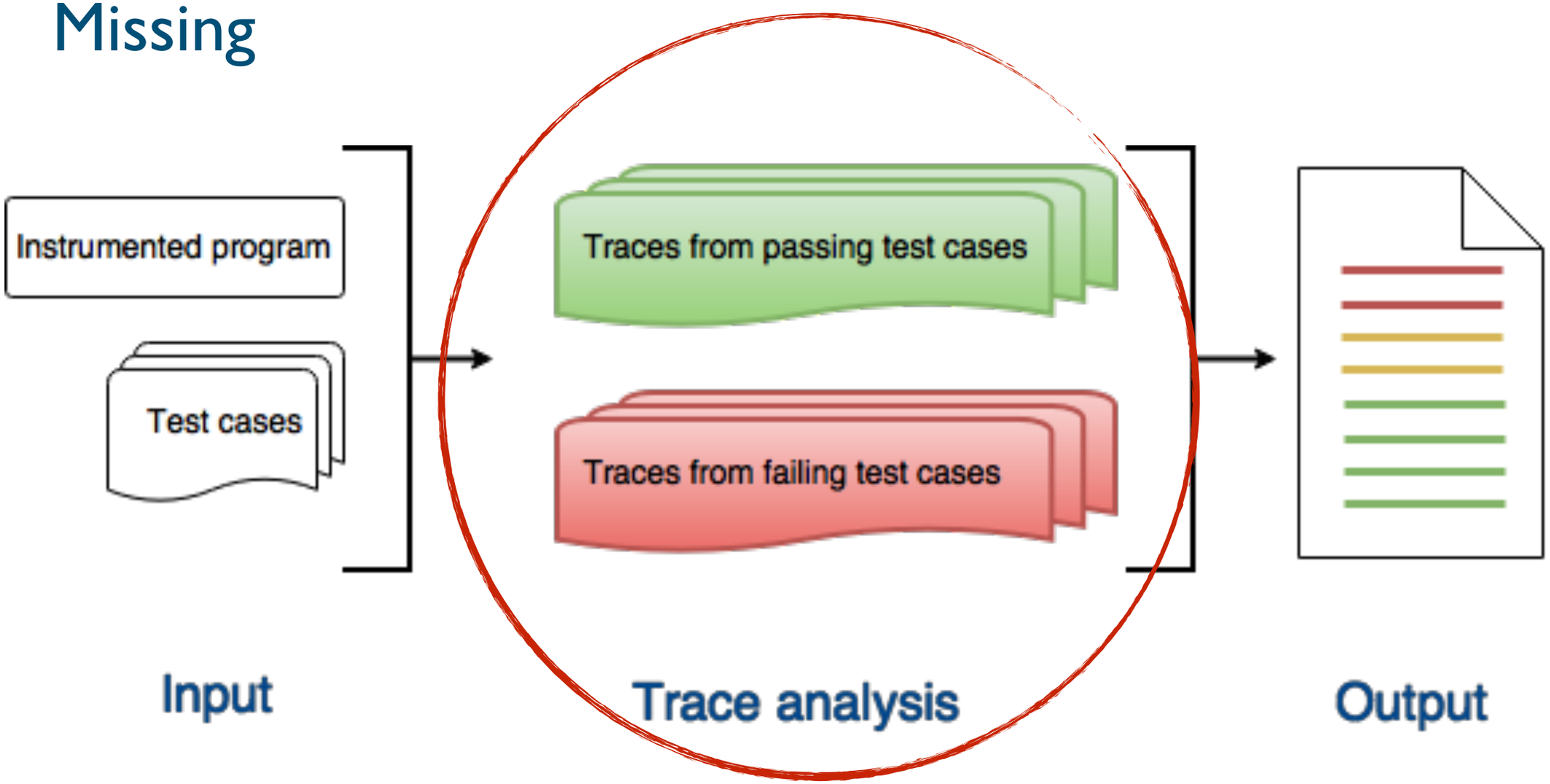



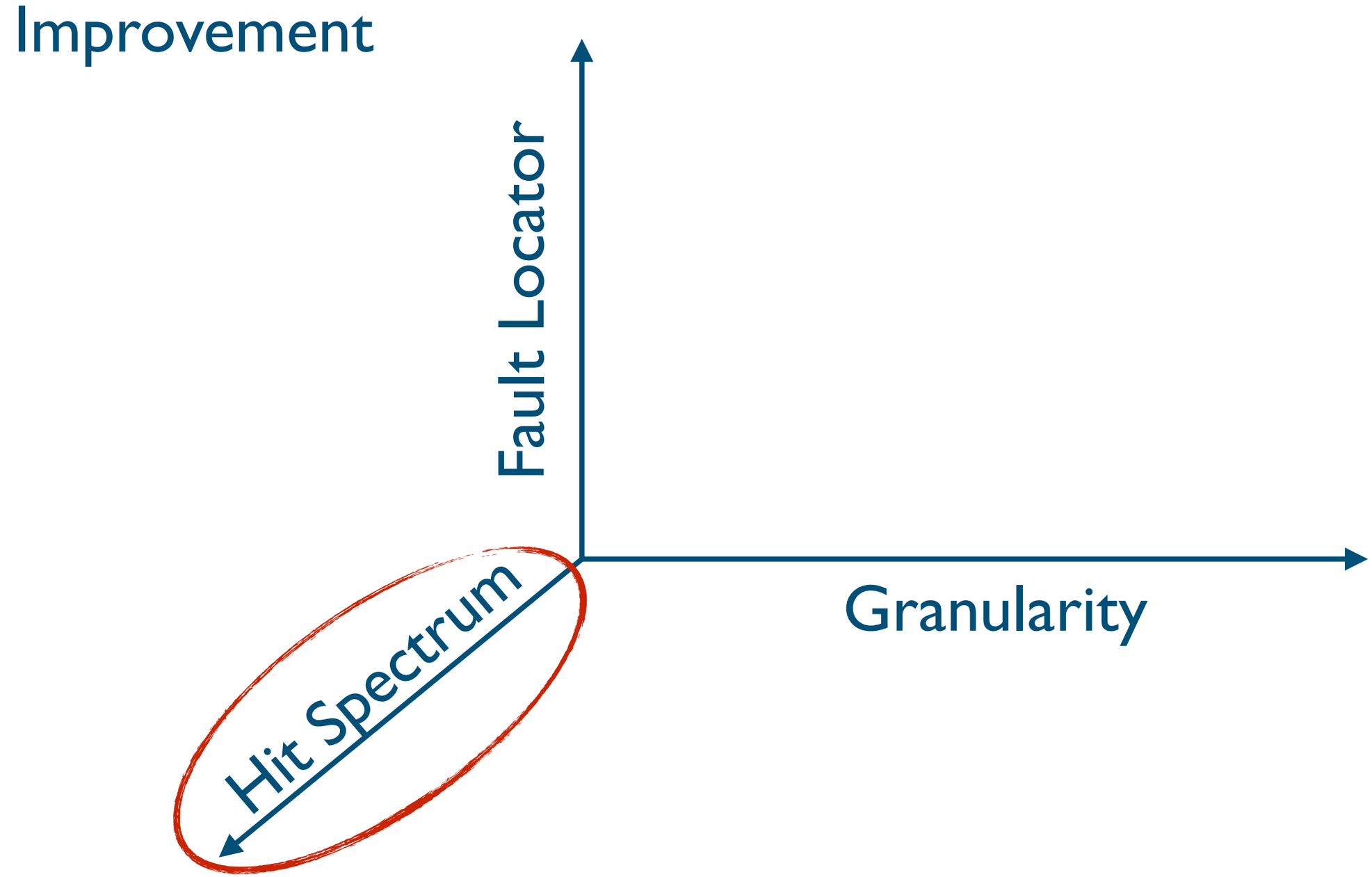
```
method1(){
   methodA()
   methodB()
   if(condition){
      return
   }
   methodC()
}
```

# Raw Spectrum Analysis

## Missing

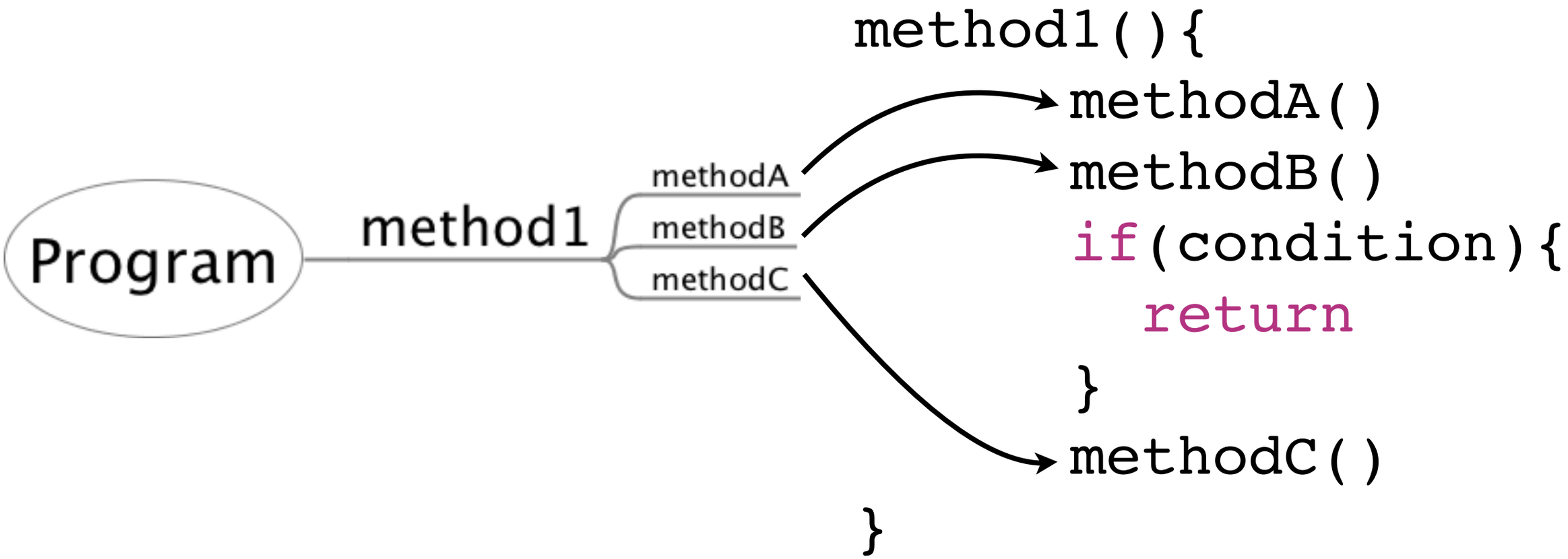


# Raw Spectrum Analysis



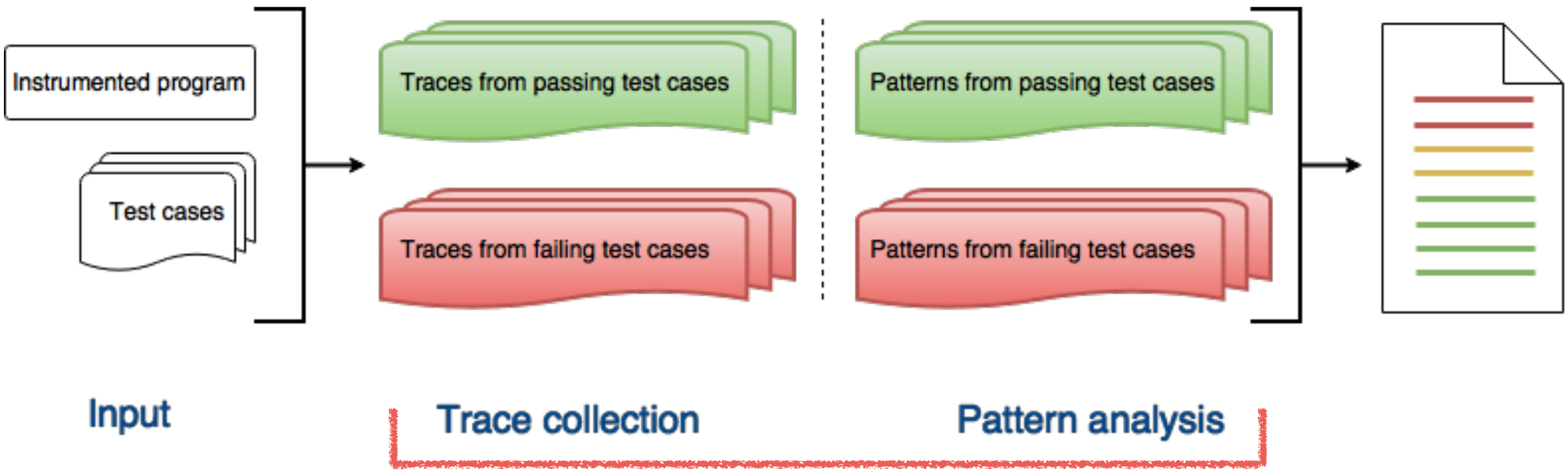
# Raw Spectrum Analysis

## Improvement



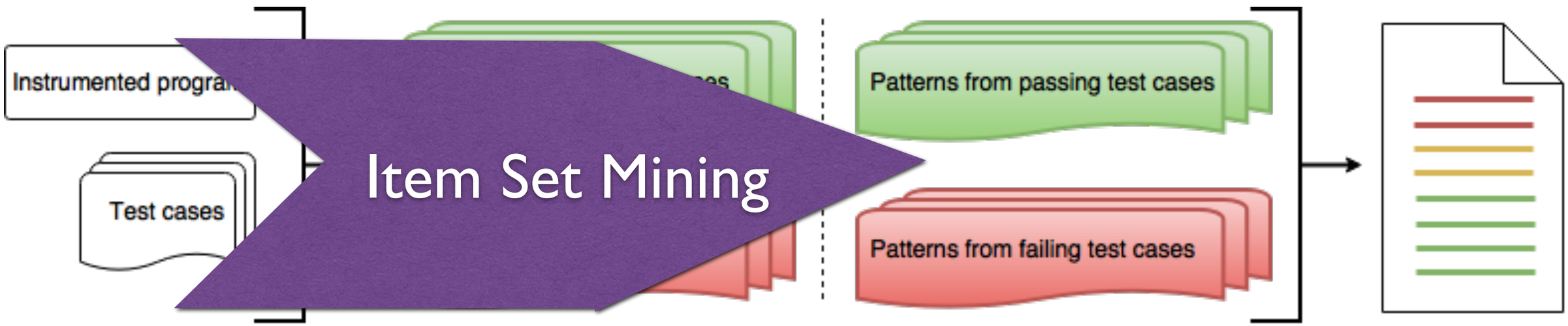
# Raw Spectrum Analysis

## Improvement



# Raw Spectrum Analysis

## Improvement



# Raw Spectrum Analysis

## Improvement

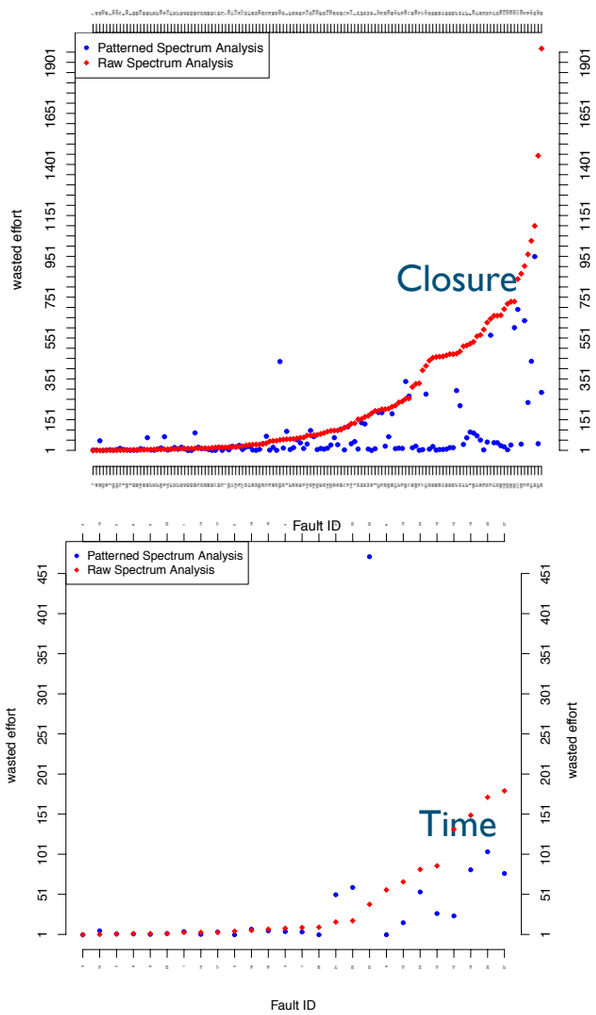
# Case Study

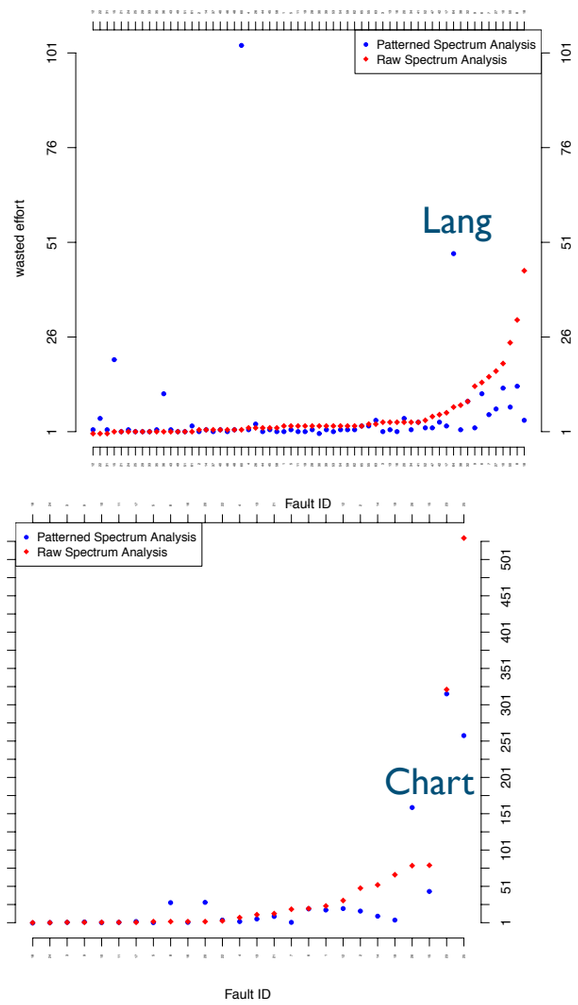

5 Open source projects

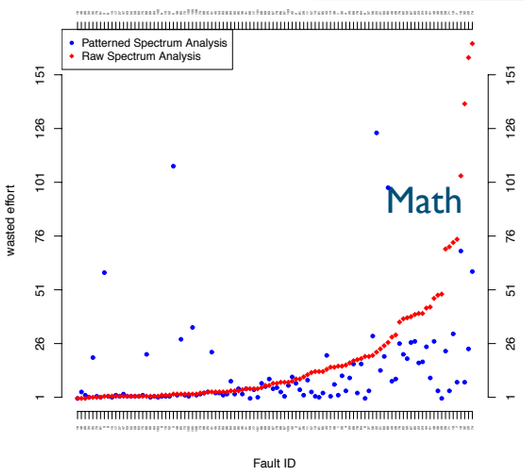
351 faults

## Evaluation metric

# Previous Study

# Patterned Spectrum Analysis

## Further improvement?

## Example: Fault in Google Closure Compiler

| | Patterned Spectrum analysis | Raw Spectrum analysis |
|---|---|---|
| Wasted effort | 85 | 532 |

# Patterned Spectrum Analysis

## Further improvement?

## Example: Fault in Google Closure Compiler

| | Patterned Spectrum analysis | Raw Spectrum analysis |
|---|---|---|
| Wasted effort | 85 | 532 |

```
Node.getLastChild()
NodeUtil.getCatchBlock(Node)
NodeUtil.hasCatchHandler(Node)
NodeUtil.hasFinally(Node)
Node.getLastChild()
tryMinimizeExits(Node,int,String)
```

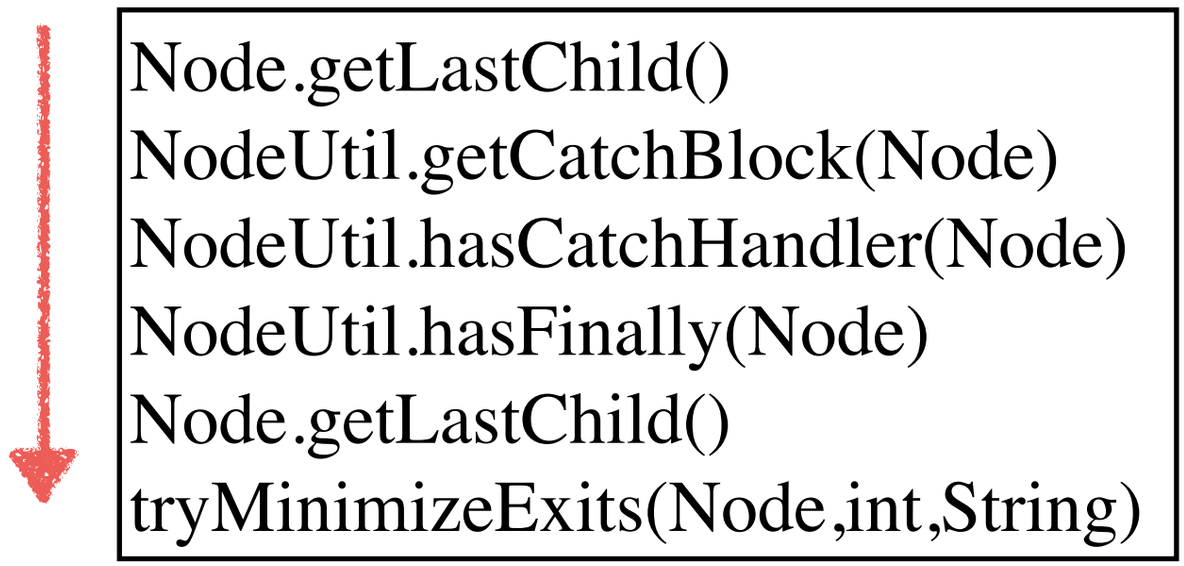
Unique pattern of a faulty method in Google Closure Compiler only in failing tests

# Patterned Spectrum Analysis

## Further improvement?

## Example: Fault in Google Closure Compiler

| | Patterned Spectrum analysis | Raw Spectrum analysis |
|---|---|---|
| Wasted effort | 85 | 532 |



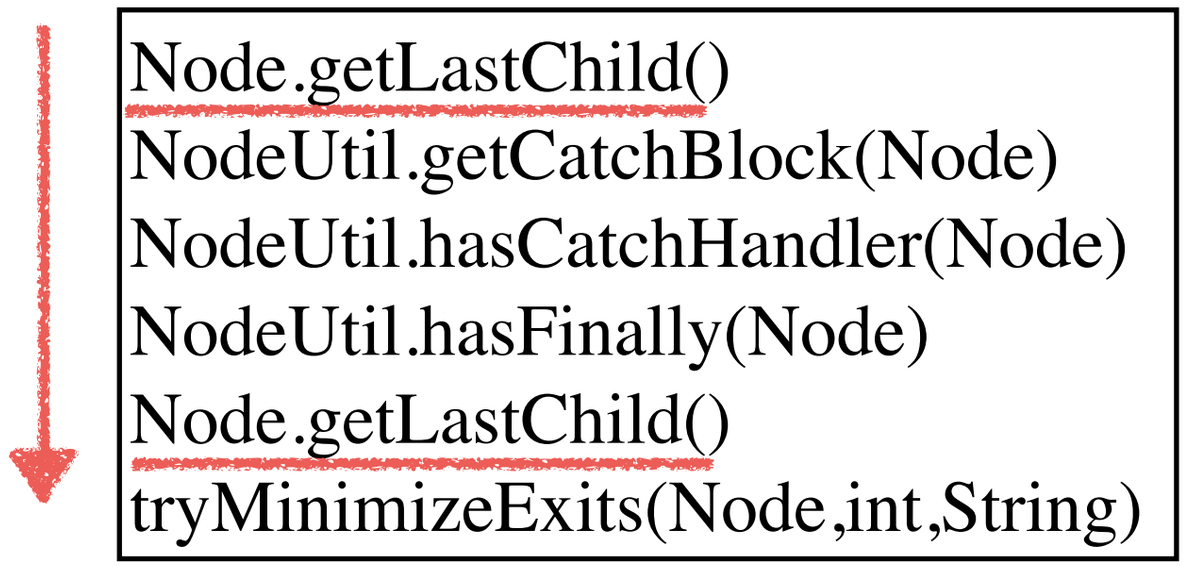

Unique pattern of a faulty method in Google Closure Compiler only in failing tests

# Patterned Spectrum Analysis

## Further improvement?

## Example: Fault in Google Closure Compiler

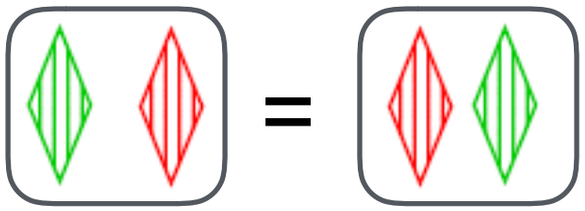
| | Patterned Spectrum analysis | Raw Spectrum analysis |
|---|---|---|
| Wasted effort | 85 | 532 |

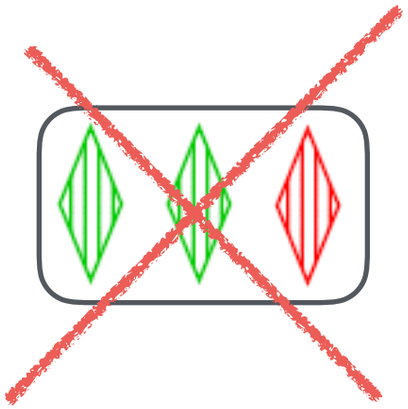



Unique pattern of a faulty method in Google Closure Compiler only in failing tests

# Patterned Spectrum Analysis

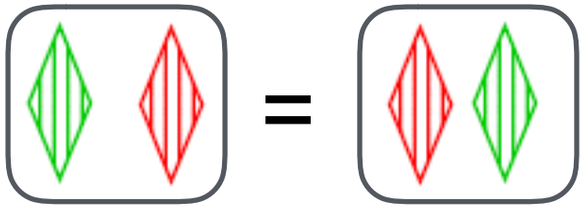
## Item Set Mining

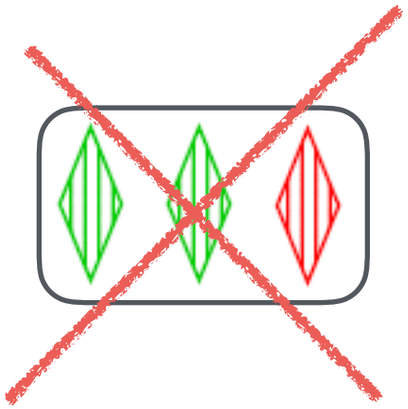




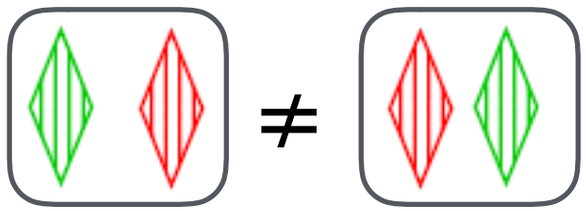
# Patterned Spectrum Analysis

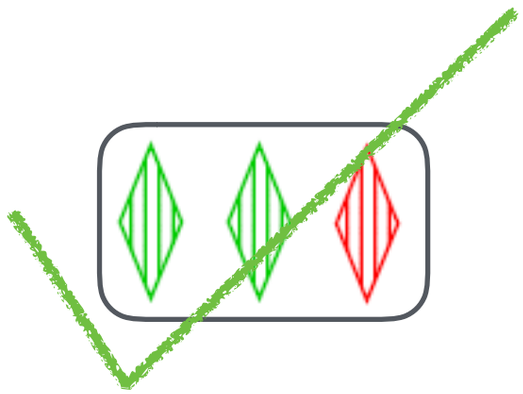
## Item Set Mining





## Sequence Mining?





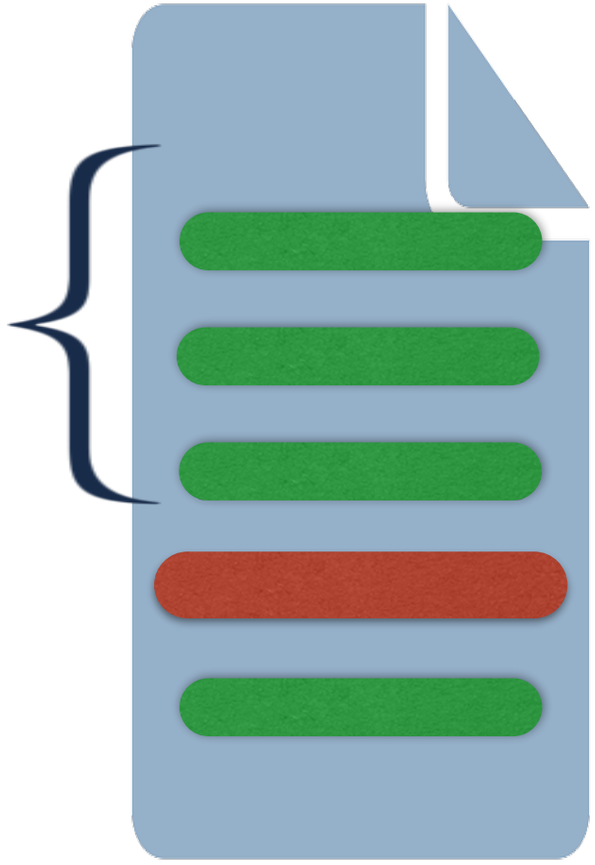
# Preliminary Study

2 Open source projects (88 faults)

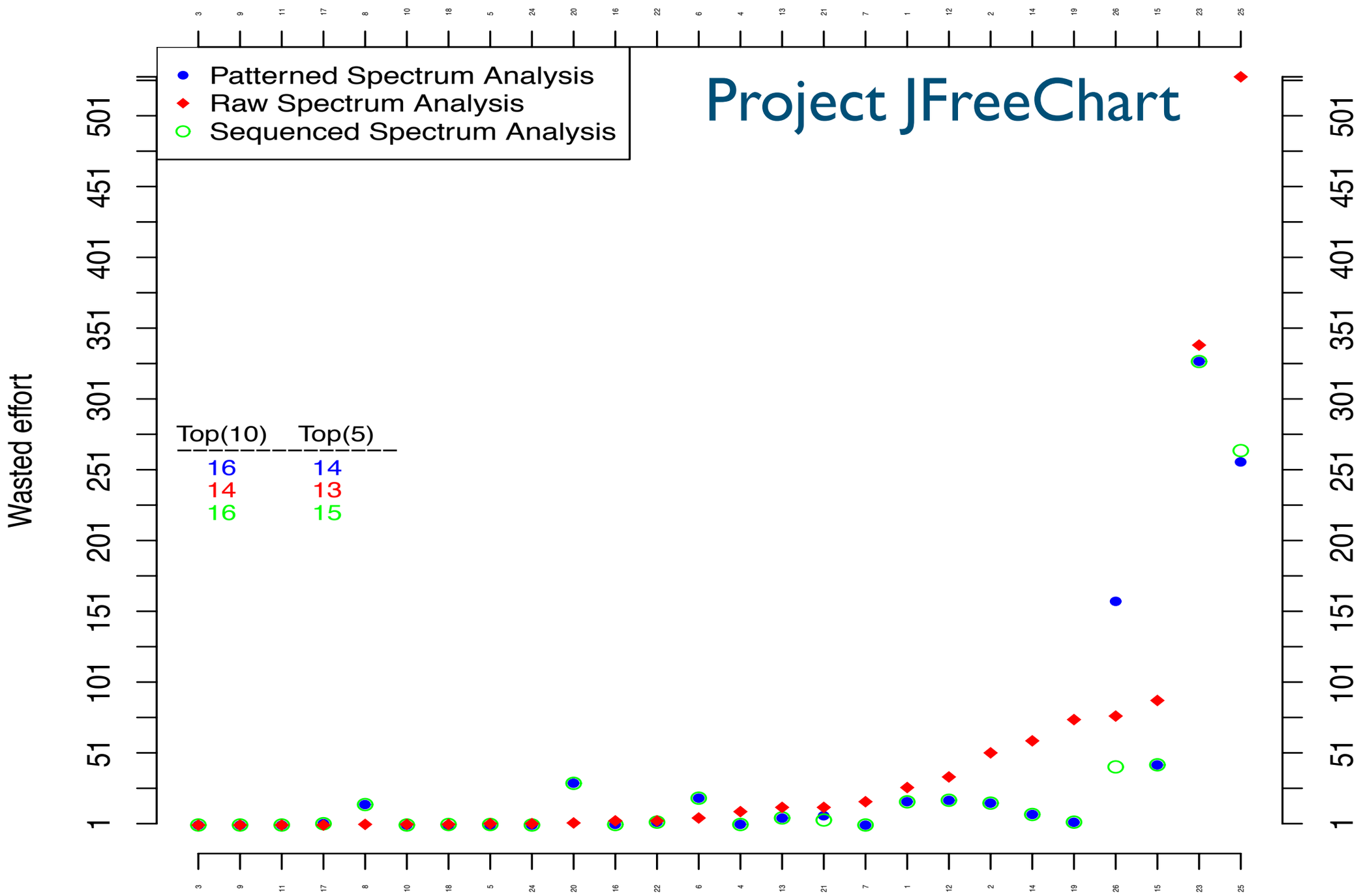
Comparison of three analyses

Evaluation metric

Wasted effort

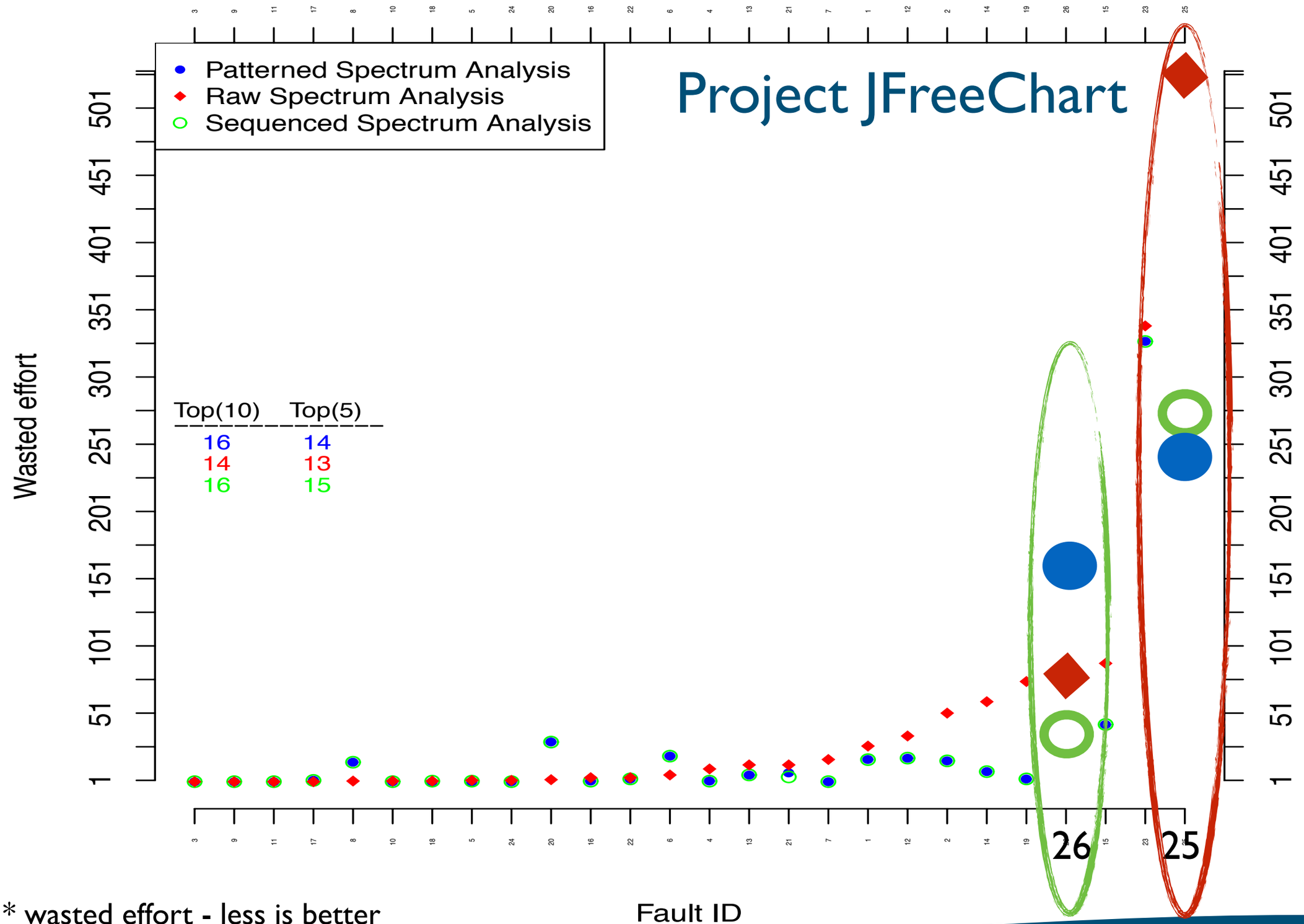


# Preliminary Results



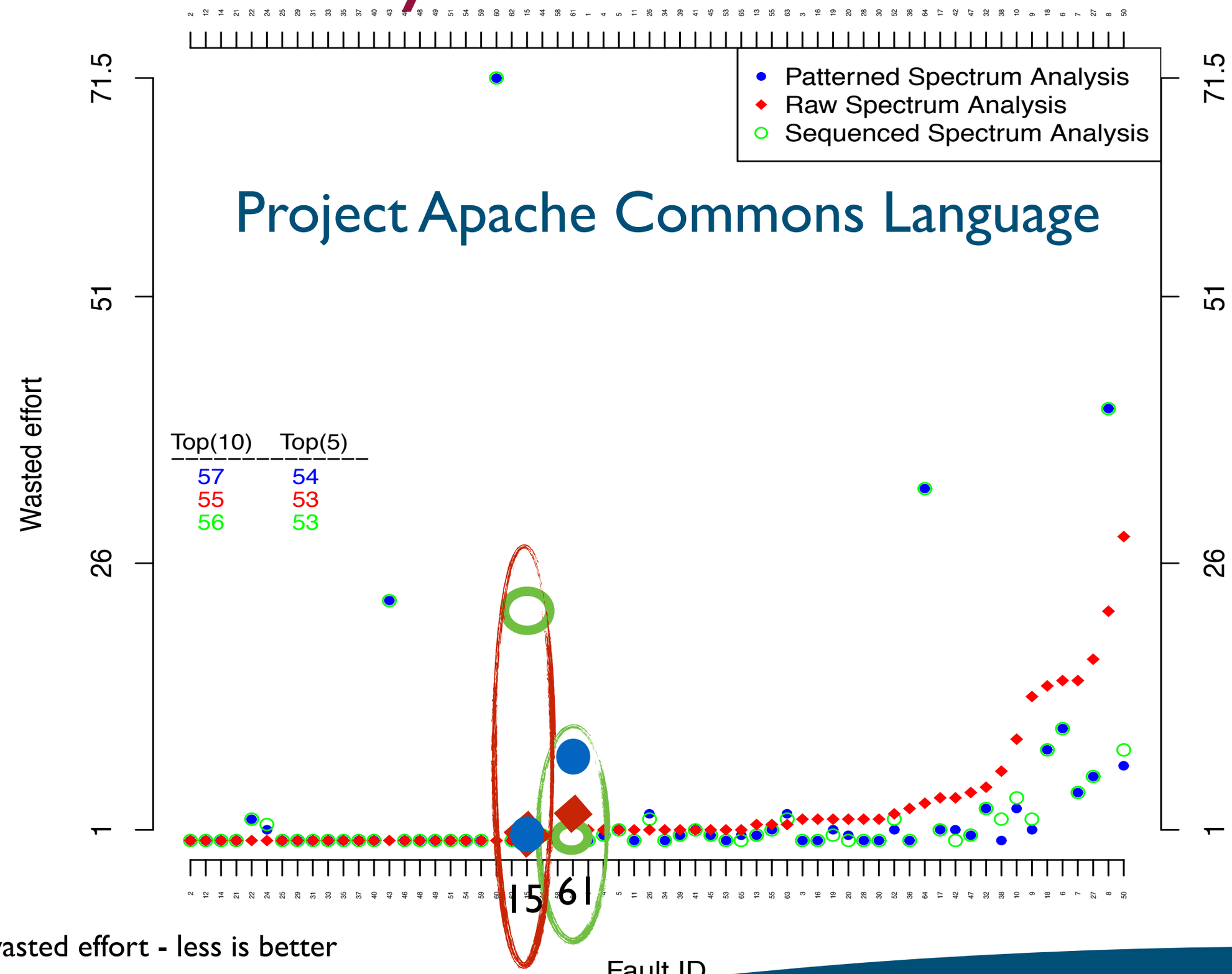

* wasted effort - less is better

# Preliminary Results




* wasted effort - less is better

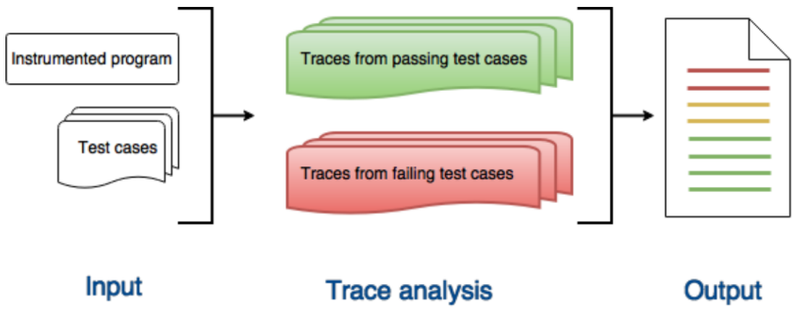
# Preliminary Results




* wasted effort - less is better

# Summary

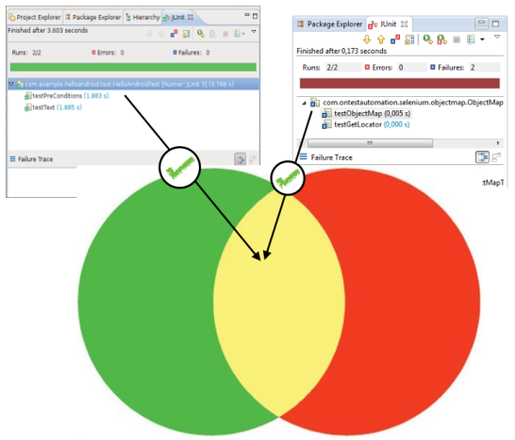
## Spectrum Based Fault Localisation






## Raw Spectrum Analysis

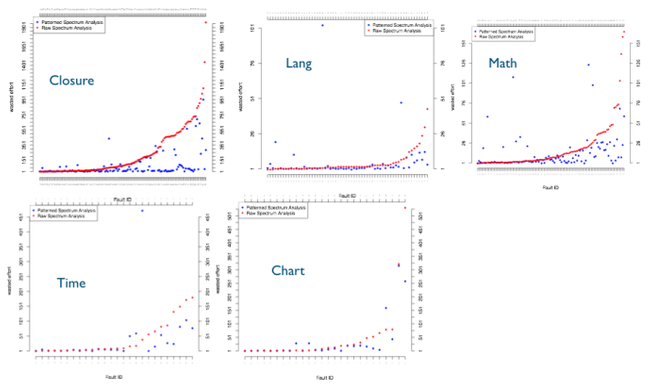
### Missing



```
public void method1(){
            ..........
            methodA()
            methodB()
            if(condition){
                return
            }
            ..........
            methodC()
            ..........
}
```

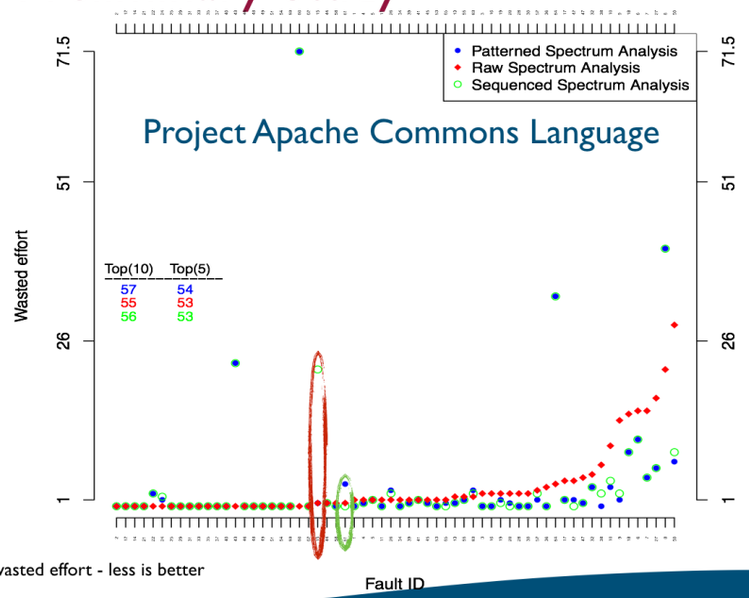


## Previous Study




## Preliminary Study




* wasted effort - less is better