On the Use of Sequence Mining within Spectrum Based Fault Localisation

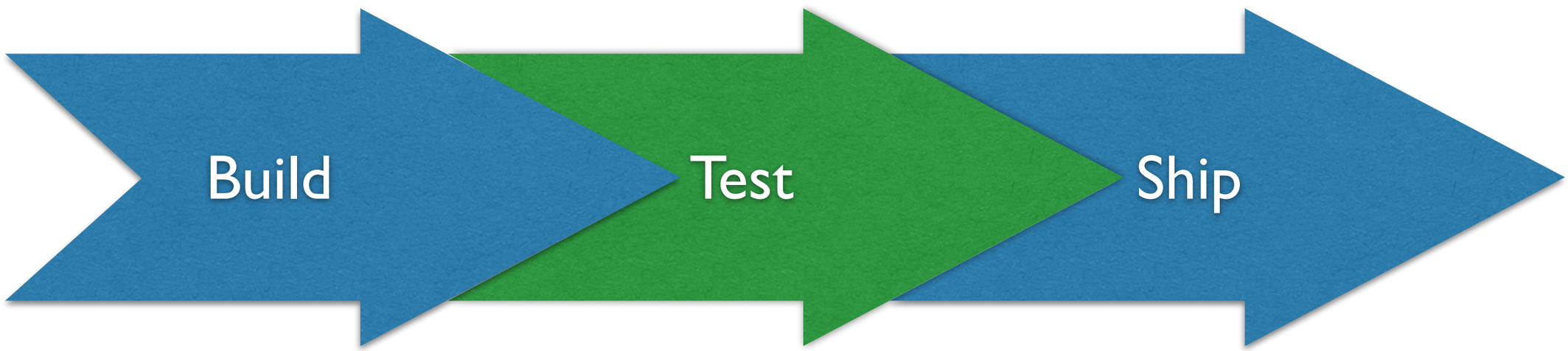
Gulsher Laghari and Serge Demeyer

SAC-SVT2018, Pau France - April 12, 2018

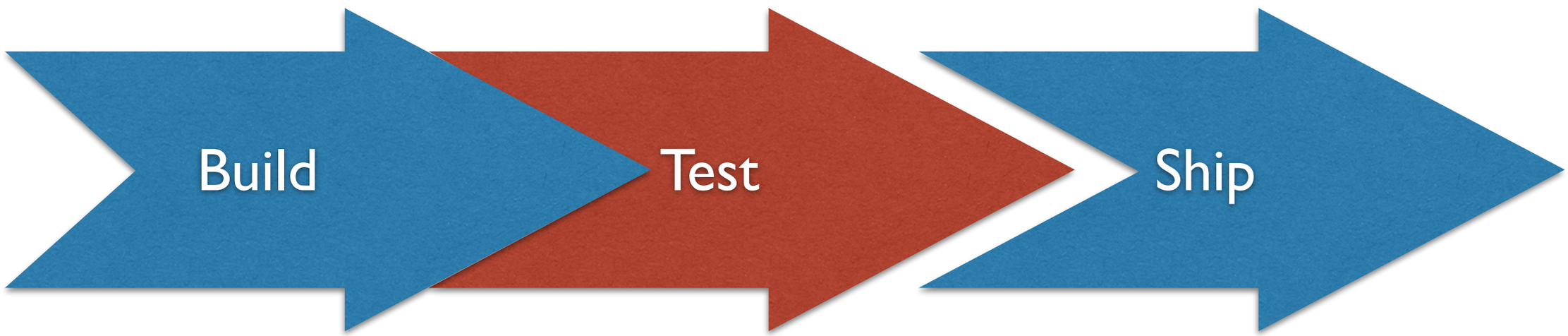
Ansymo
Antwerp Systems & Software Modelling
University of Antwerp

Universiteit Antwerpen

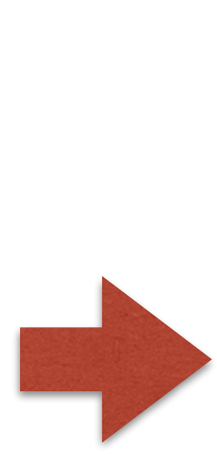
Overview



Overview



Fault Localisation

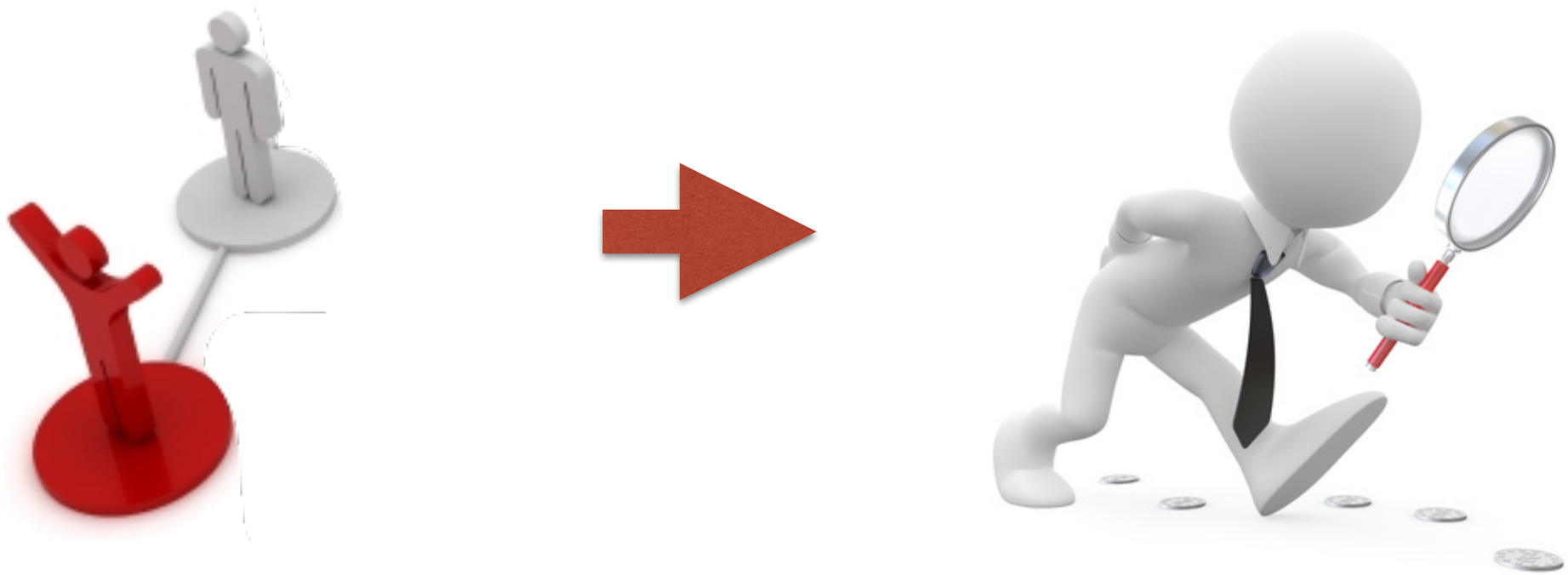


Locate Understand Fix

Fault Localisation an important step in debugging process

Fault Localisation

Test to code mapping - 1:1



Fault Localisation

Test to code mapping - 1:N

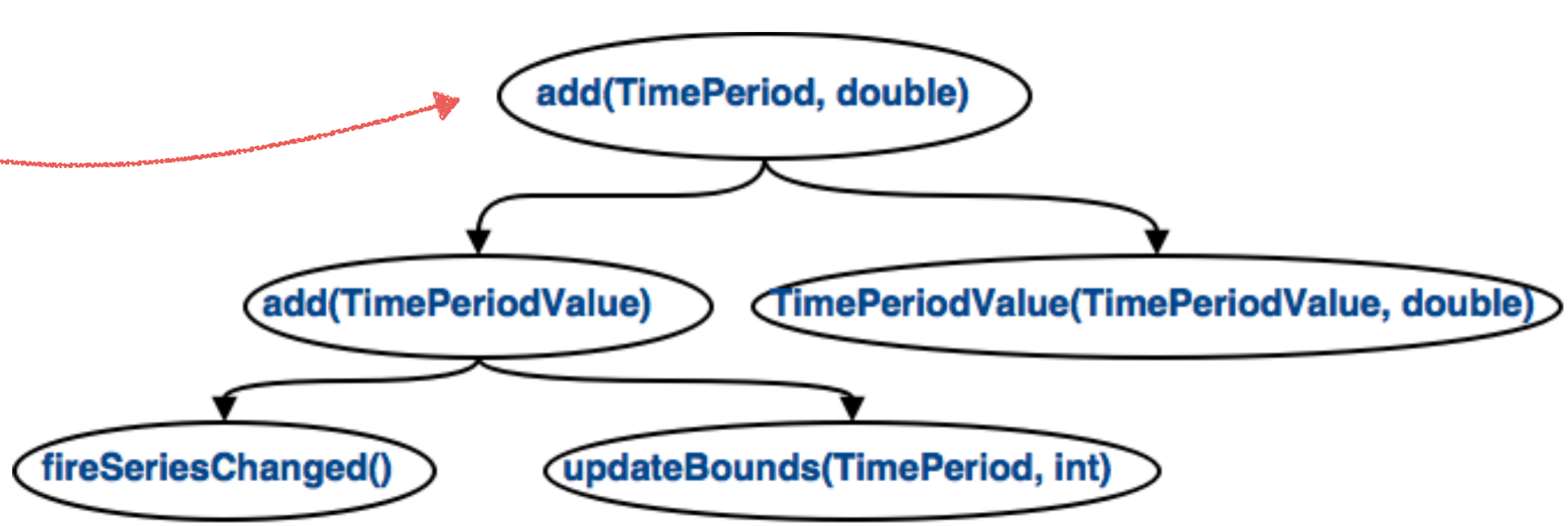
Fault Localisation

```
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

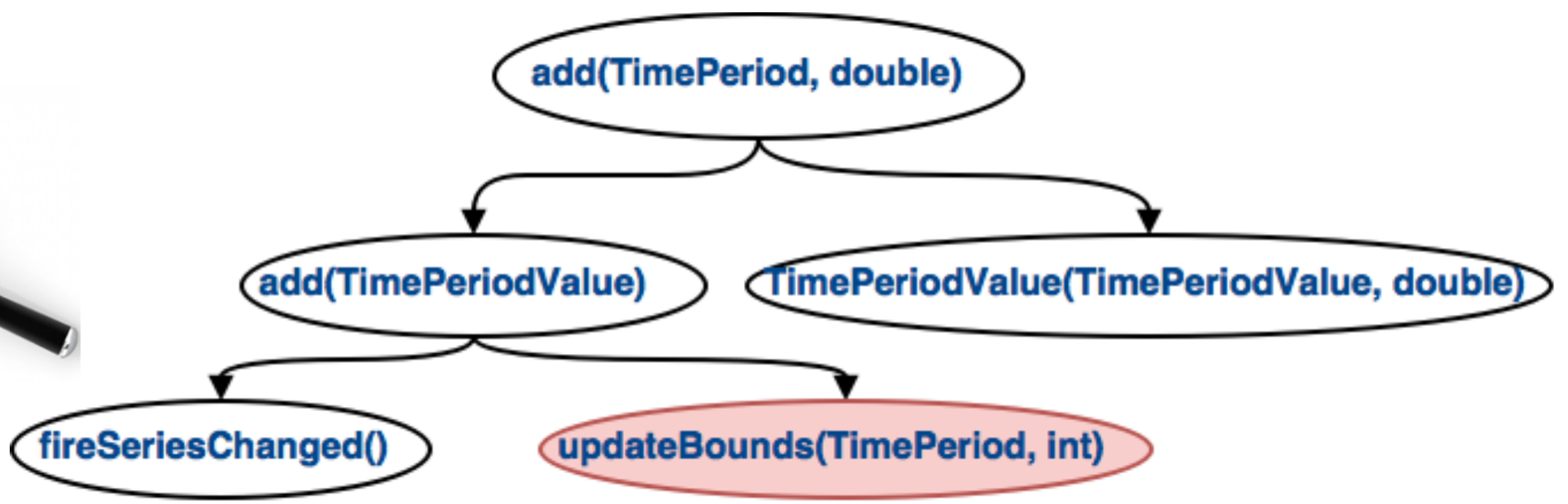
Fault Localisation

```
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

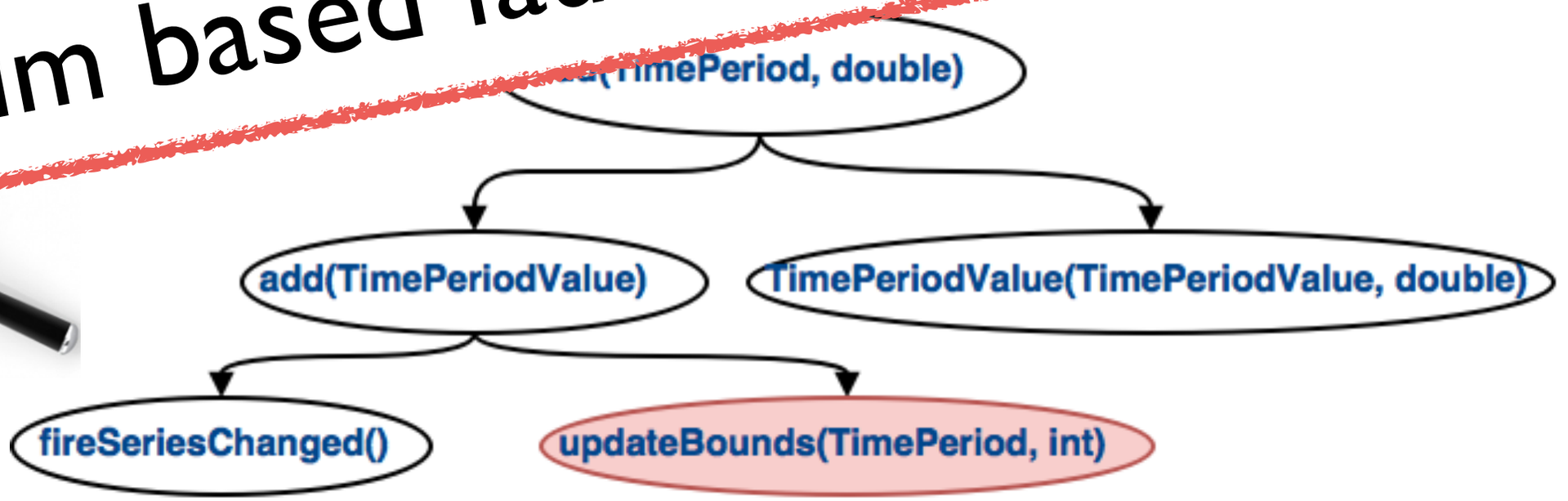
```
public void testGetMaxMiddleIndex() {
        TimePeriodValues s = new TimePeriodValues("Test");
        assertEquals(-1, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(100L, 200L), 1.0);
        assertEquals(0, s.getMaxMiddleIndex());
        s.add(new SimpleTimePeriod(300L, 400L), 2.0);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



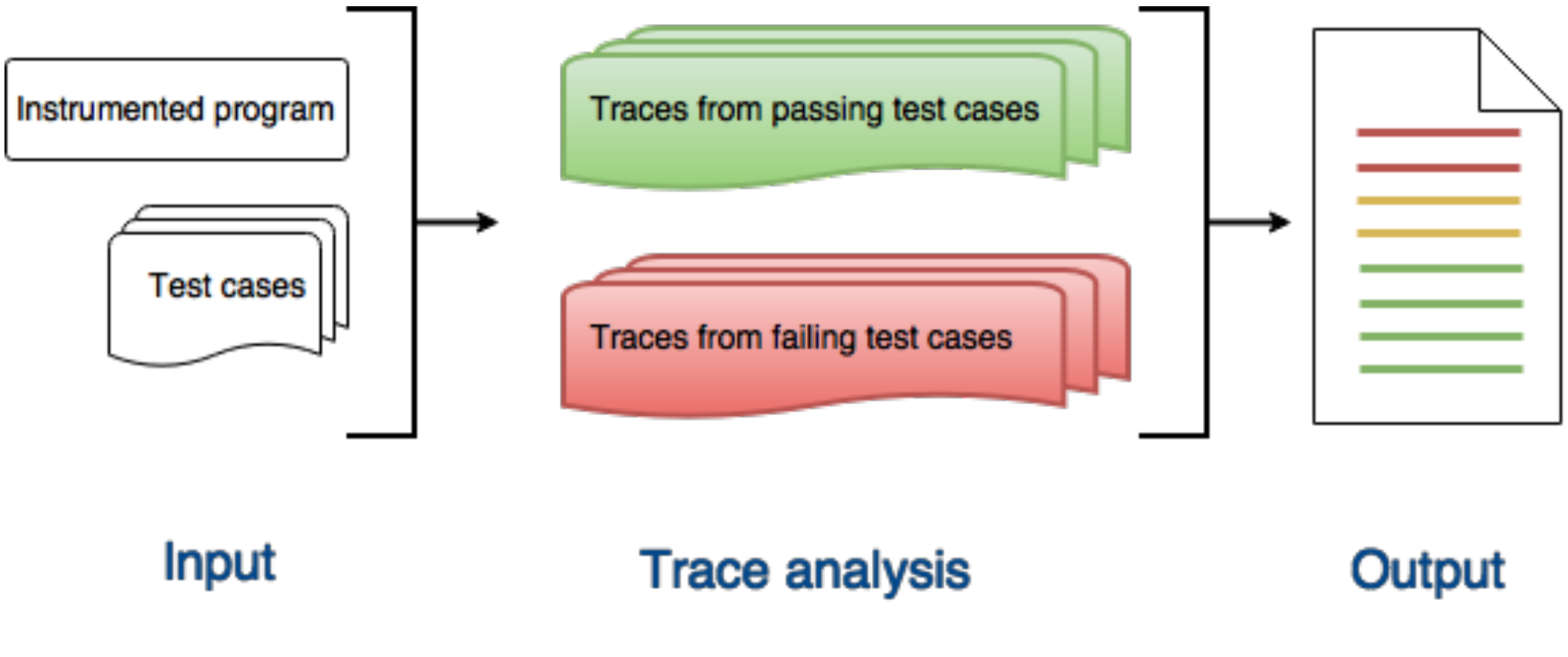
Fault Localisation

```
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());
    }
```

Spectrum based fault localisation

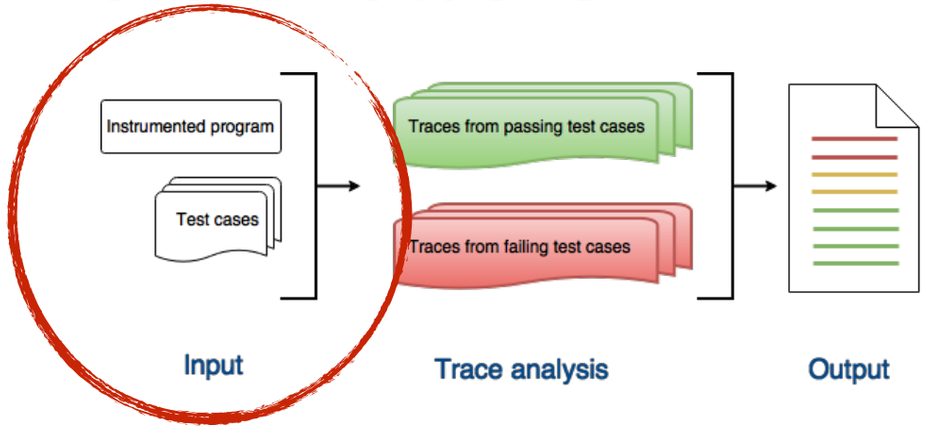


Spectrum Based Fault Localisation



Spectrum Based Fault Localisation

Granularity

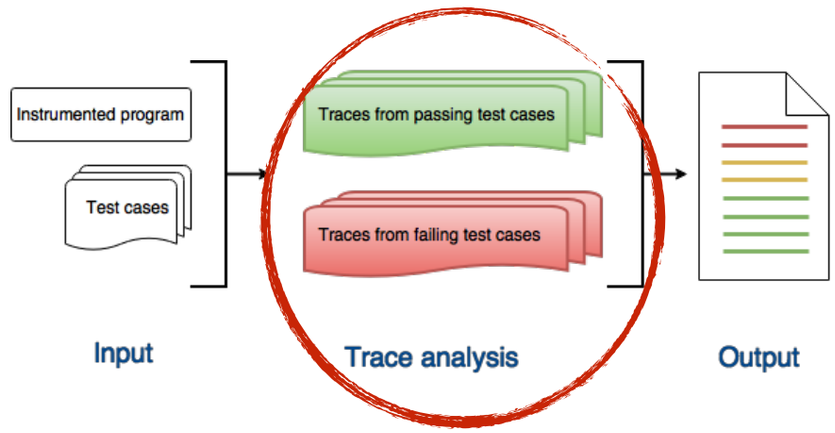


Statement | Block | Method | Class

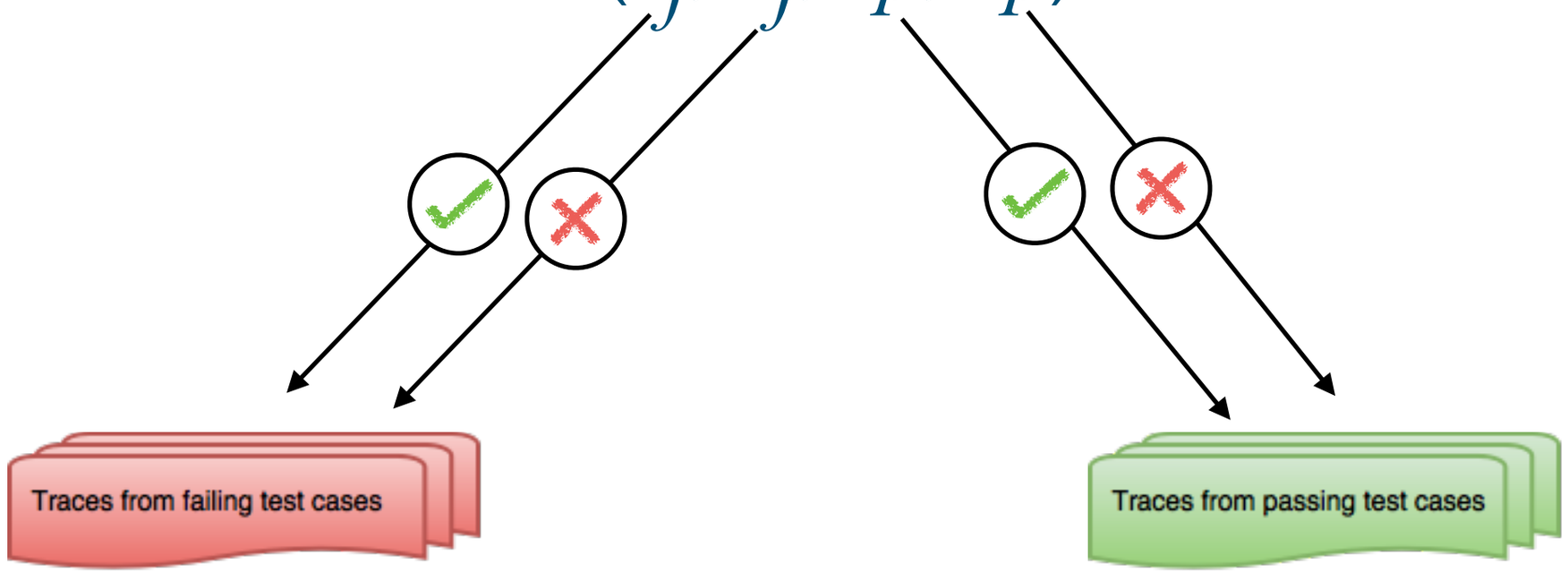
XS | S | M | L

Spectrum Based Fault Localisation

Hit Spectrum



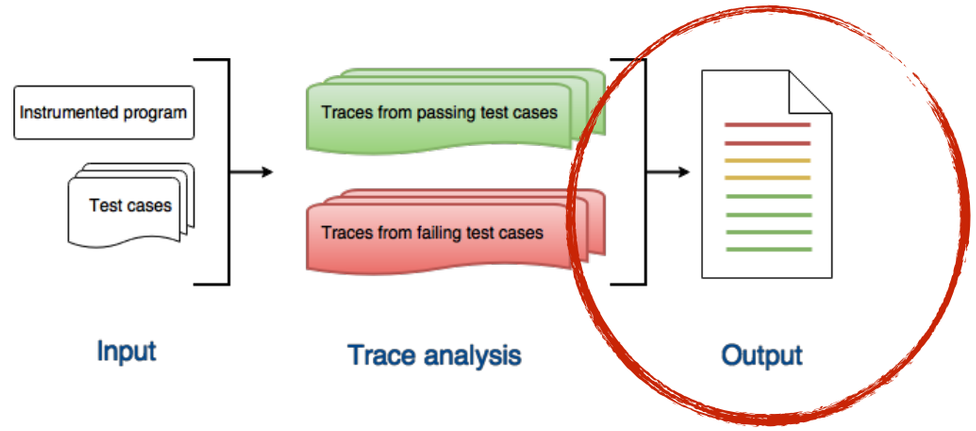
Method = (e_f, n_f, e_p, n_p)



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

Spectrum Based Fault Localisation

Fault Locator

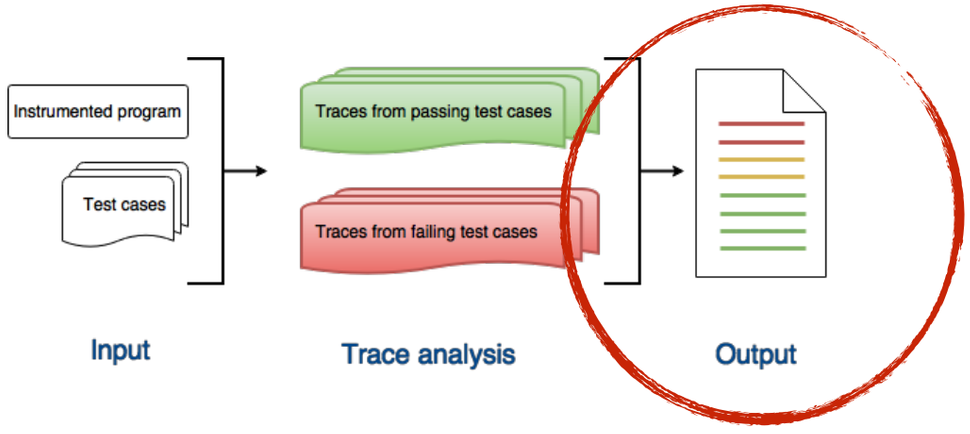


Method = (e_f, n_f, e_p, n_p) Method = Suspiciousness [0,1]

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

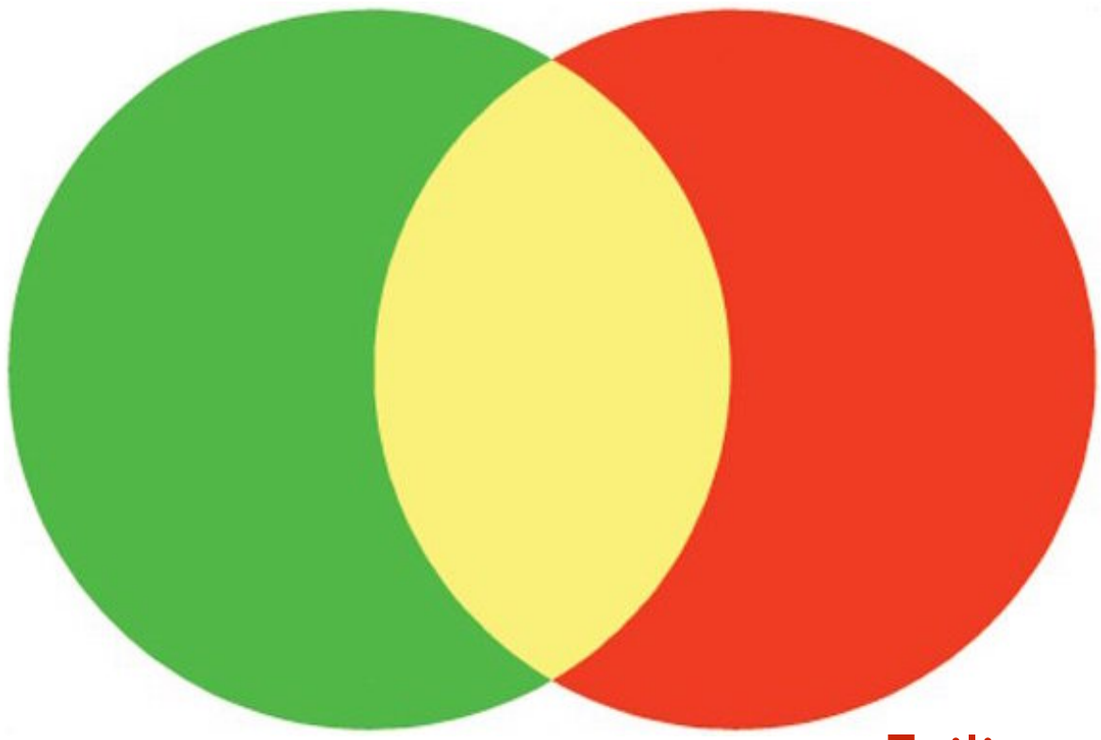
Spectrum Based Fault Localisation

Fault Locator



Not suspicious at all → Highly suspicious

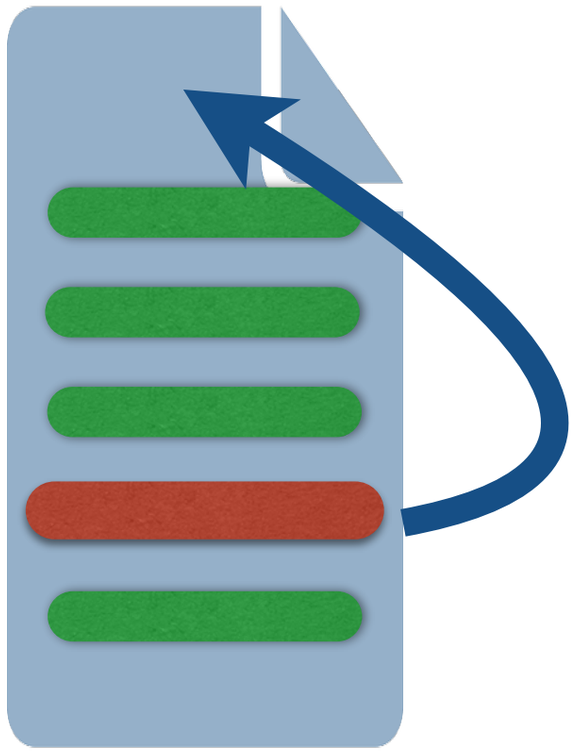
0 — 1



Passing test cases

Failing test cases

Spectrum Based Fault Localisation



Spectrum Based Fault Localisation

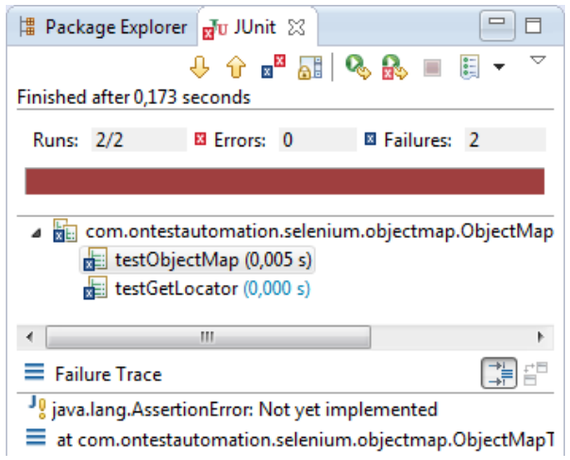
Raw Spectrum Analysis

Missing

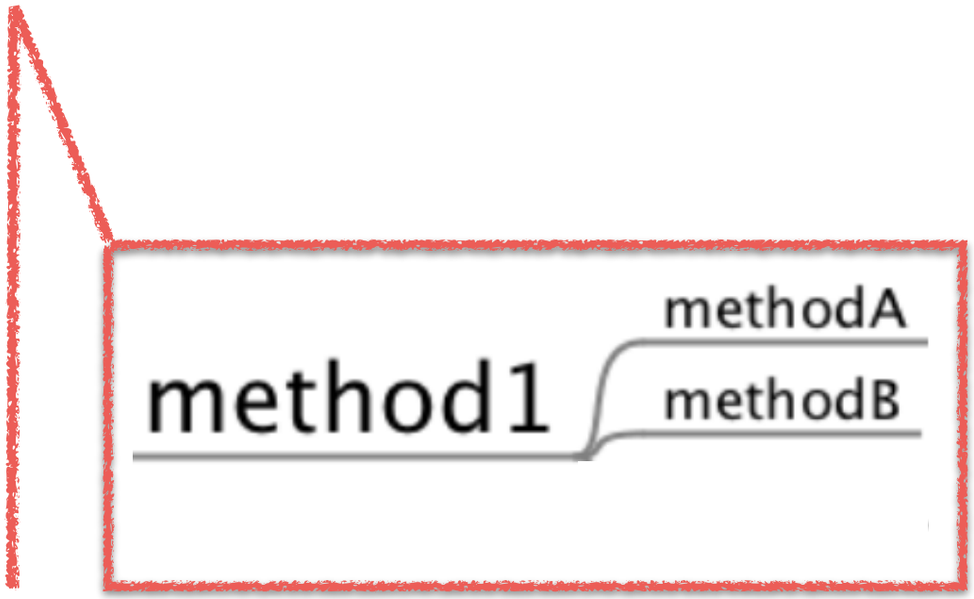
```
method1(){
        methodA()
        methodB()
        if(condition){
          return
        }
        methodC()
}
```

Raw Spectrum Analysis

Missing

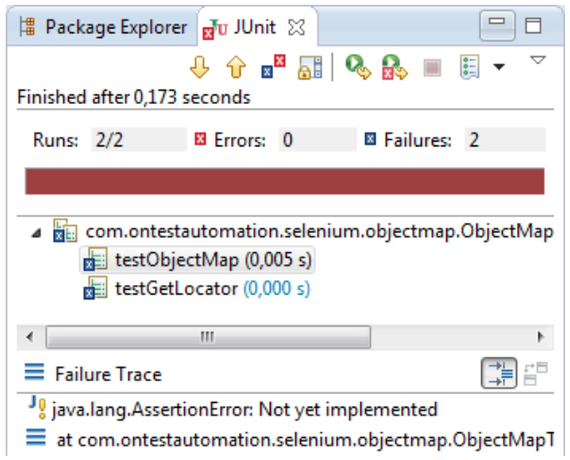



```
method1(){
   methodA()
   methodB()
   if(condition){
      return
   }
   methodC()
}
```

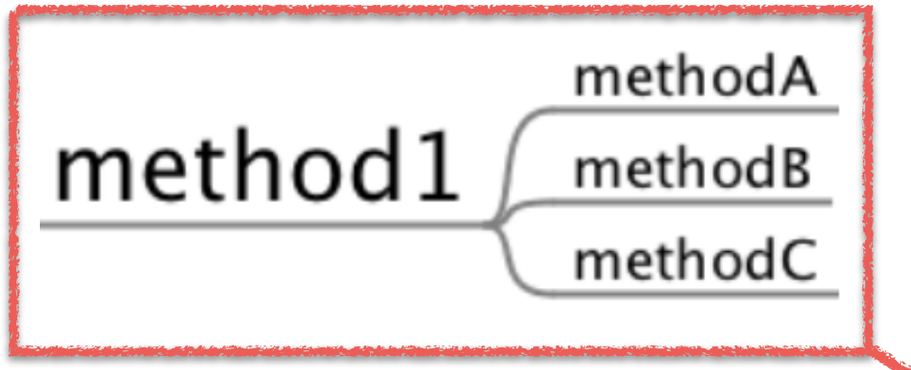


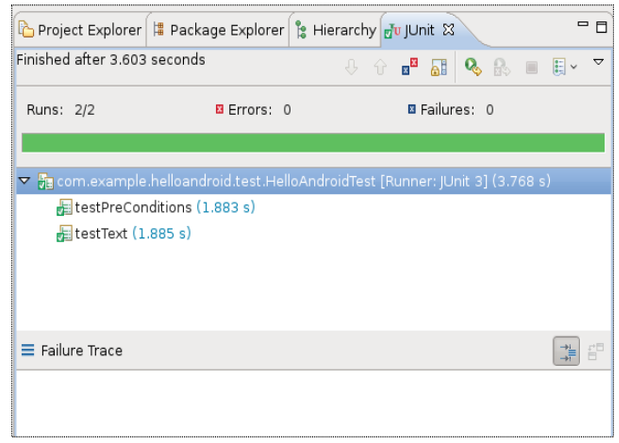
Raw Spectrum Analysis

Missing

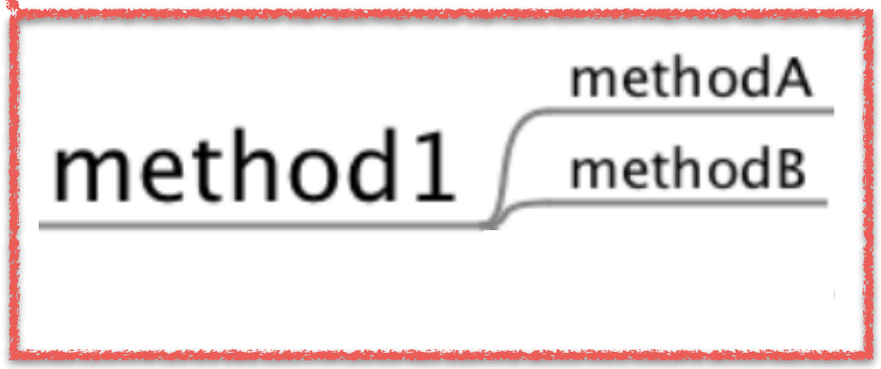



```
method1(){
   methodA()
   methodB()
   if(condition){
      return
   }
   methodC()
}
```



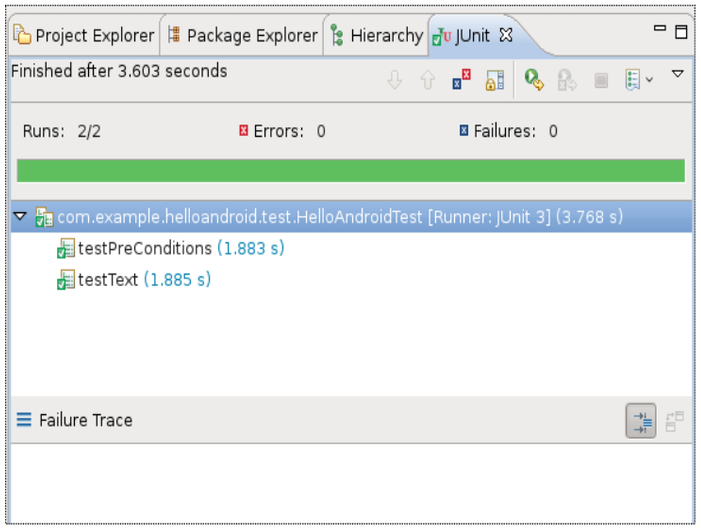


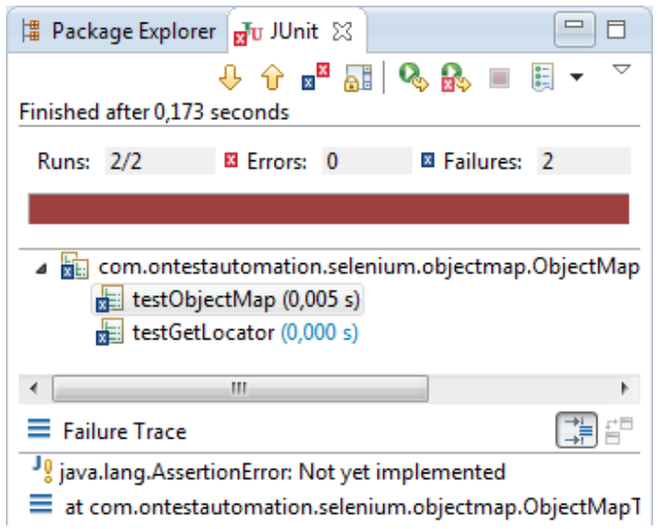

```
method1(){
   methodA()
   methodB()
   if(condition){
      return
   }
   methodC()
}
```

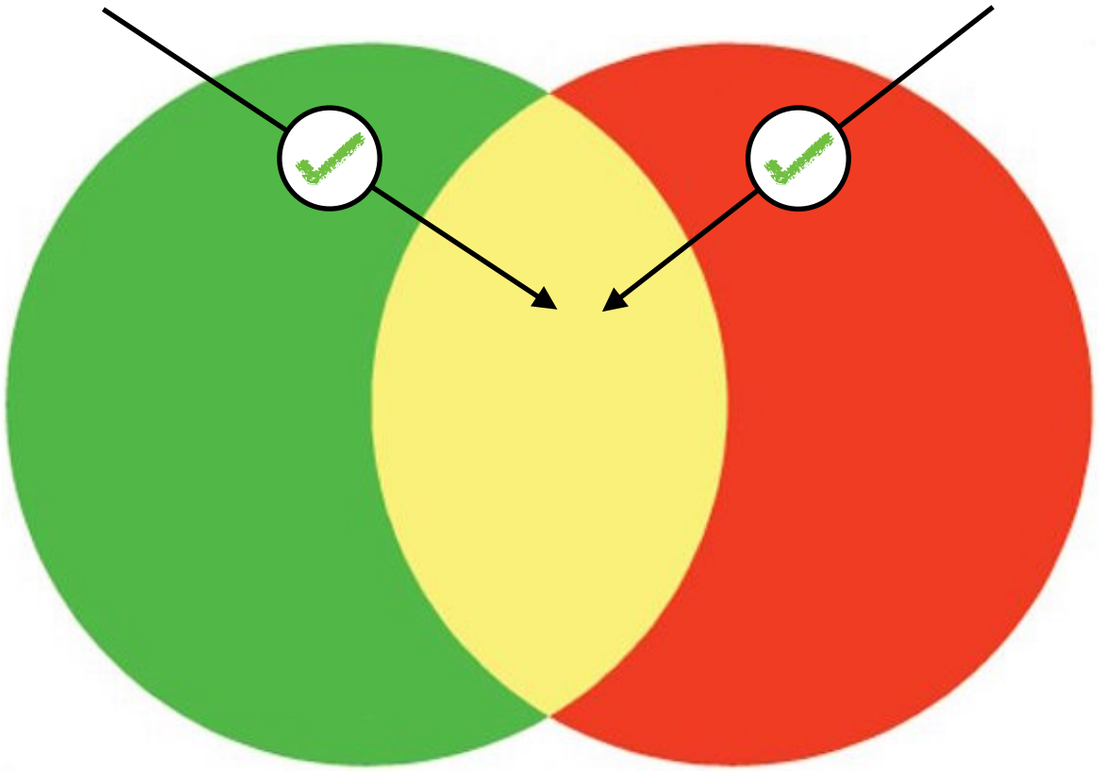


Raw Spectrum Analysis

Missing



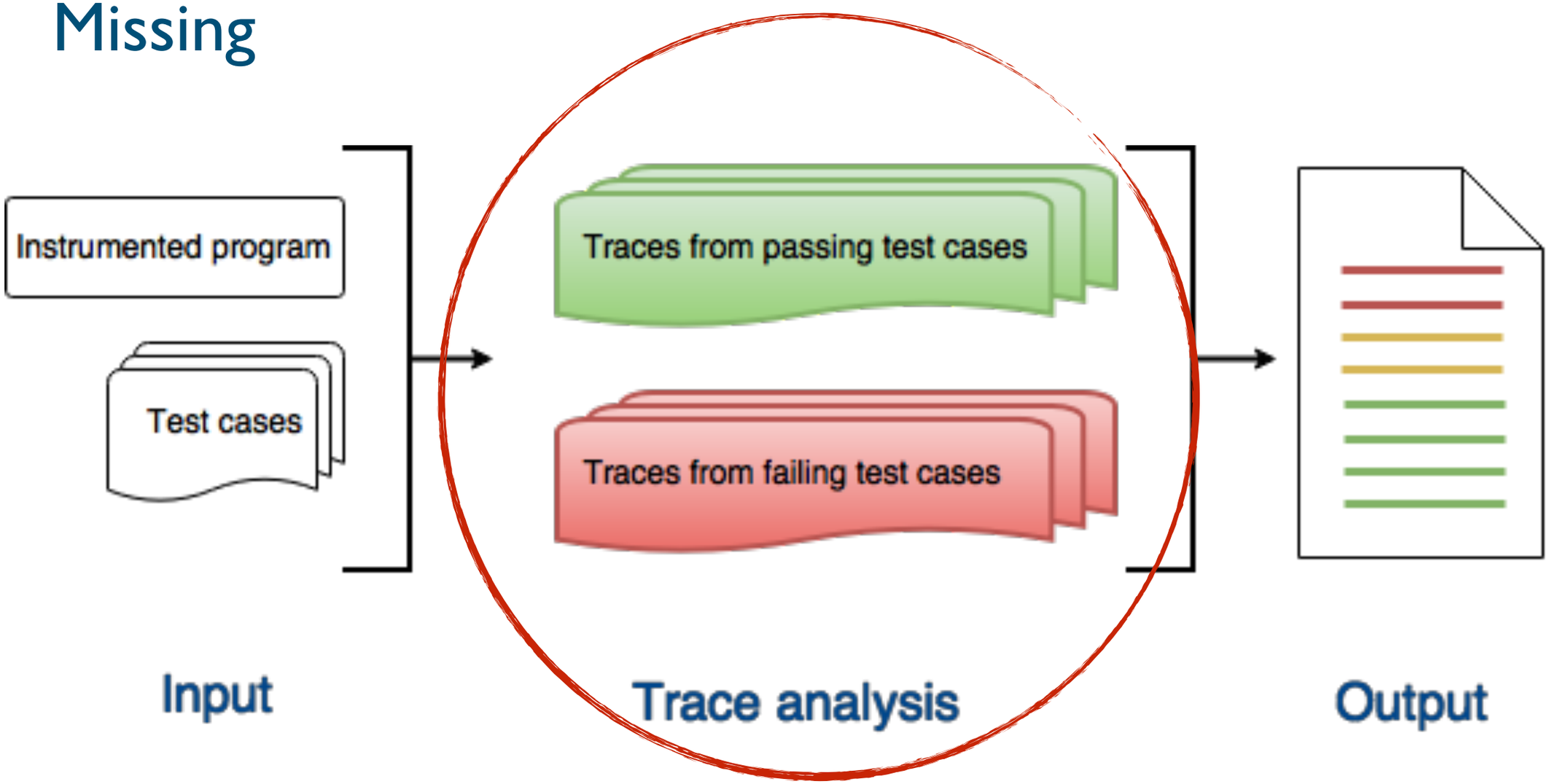




```
method1(){
   methodA()
   methodB()
   if(condition){
      return
   }
   methodC()
}
```

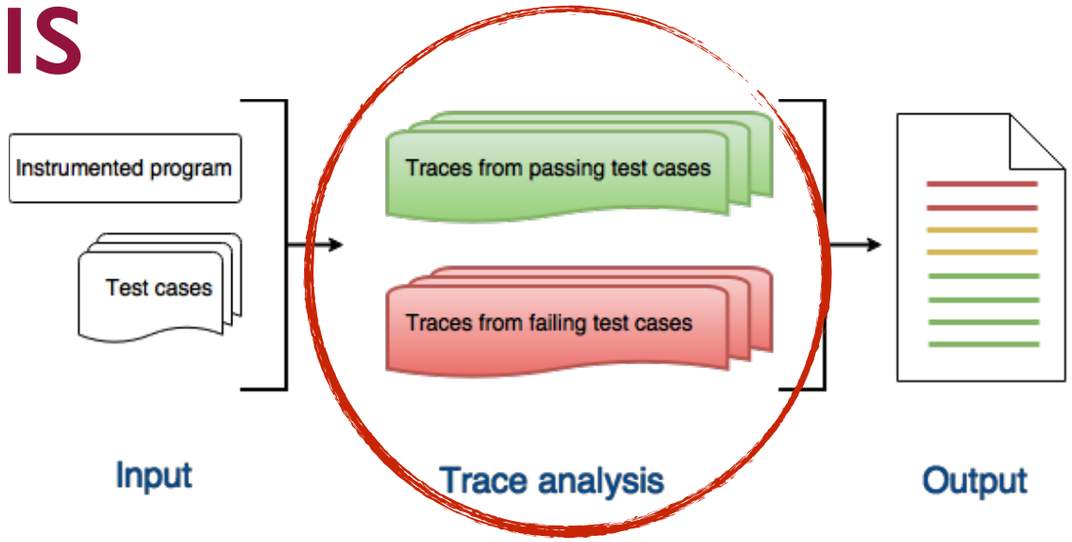
Raw Spectrum Analysis

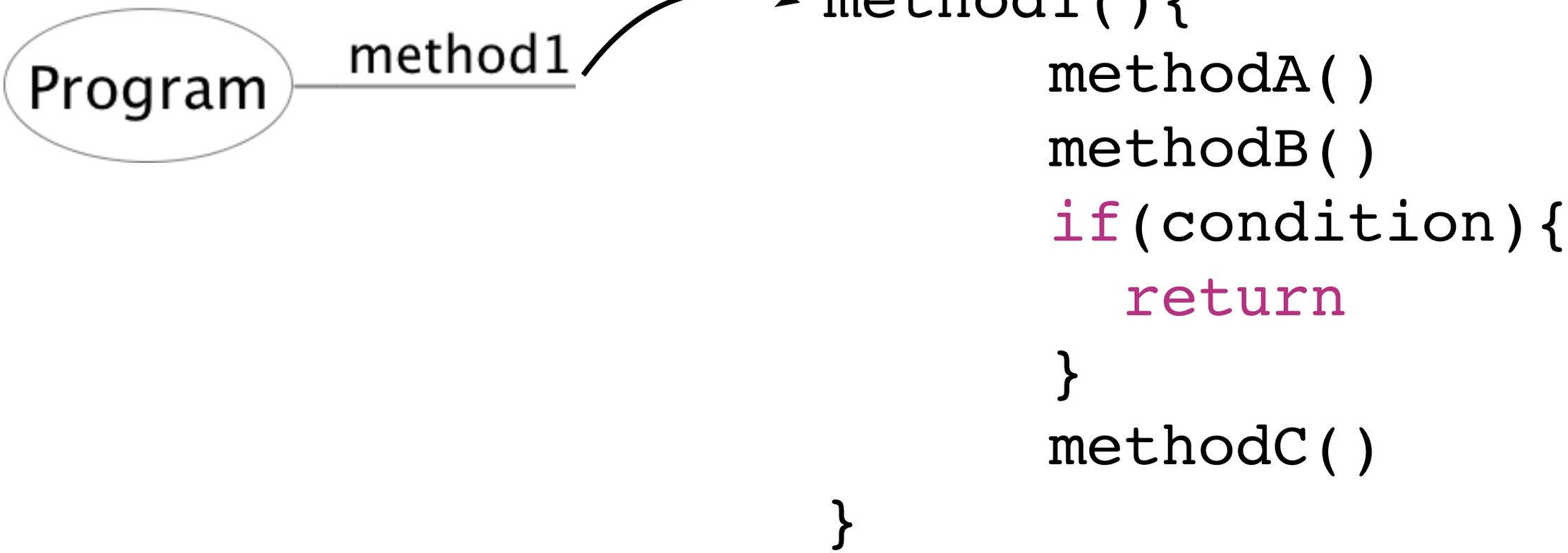
Missing



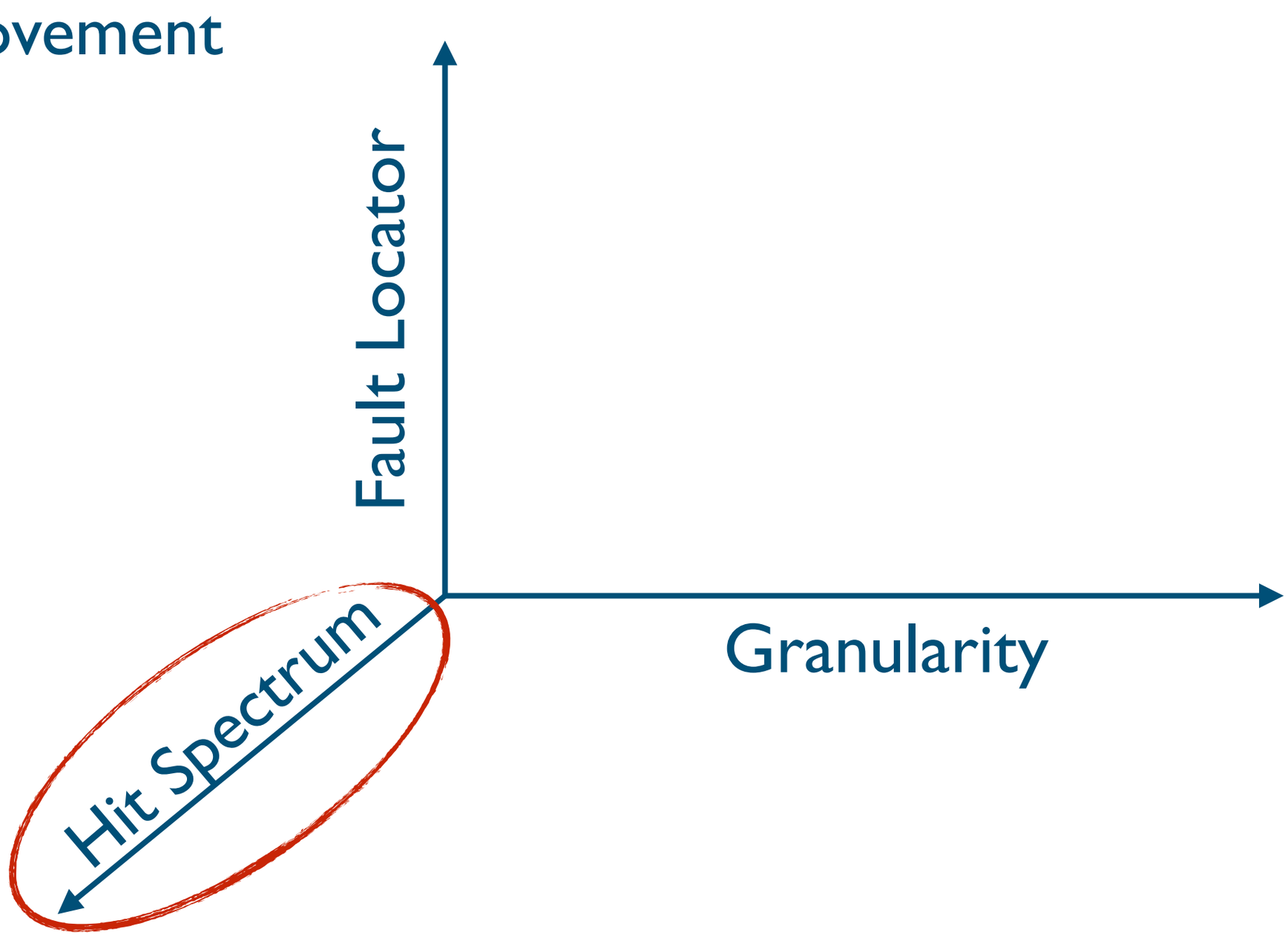
Raw Spectrum Analysis

Missing



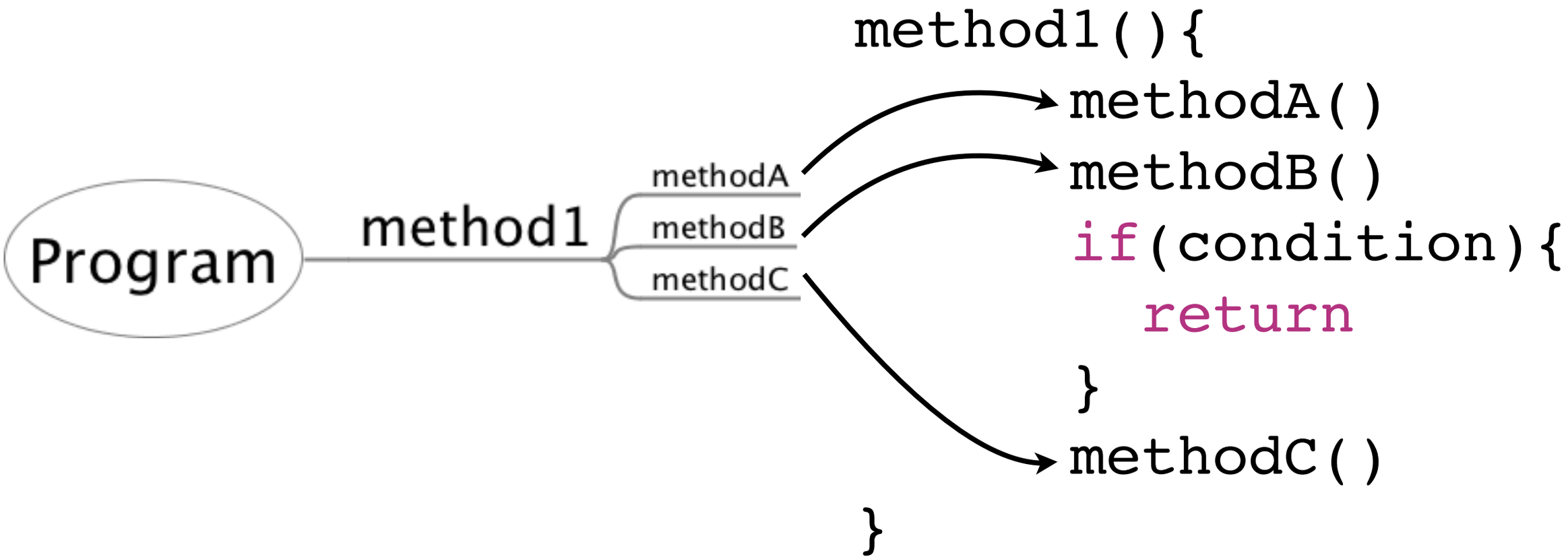


Raw Spectrum Analysis



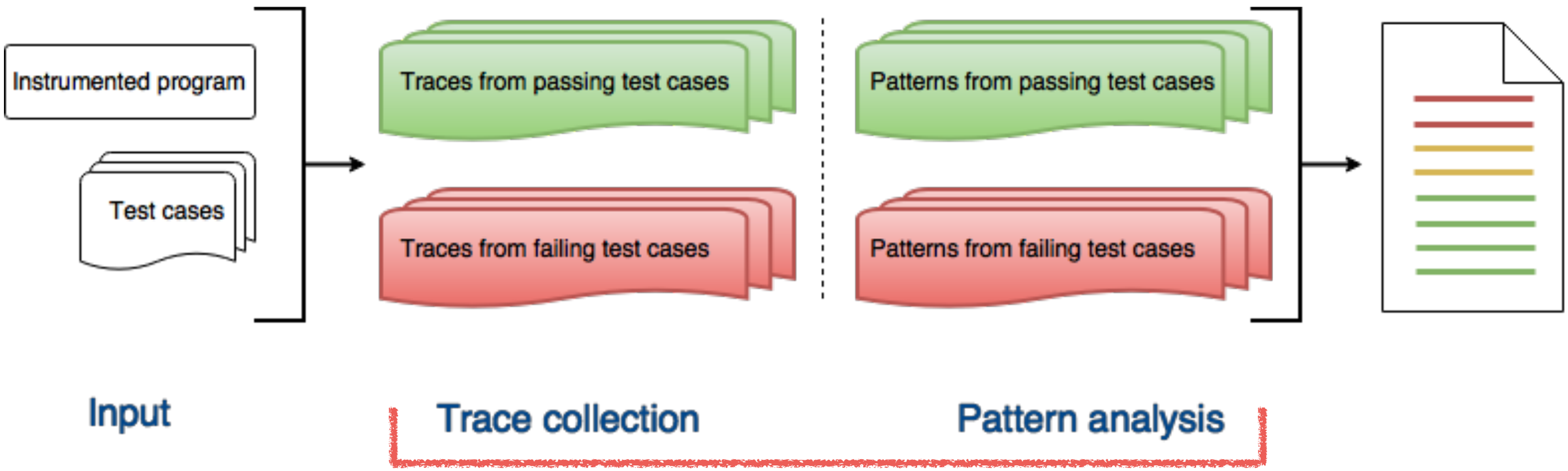
Raw Spectrum Analysis

Improvement



Raw Spectrum Analysis

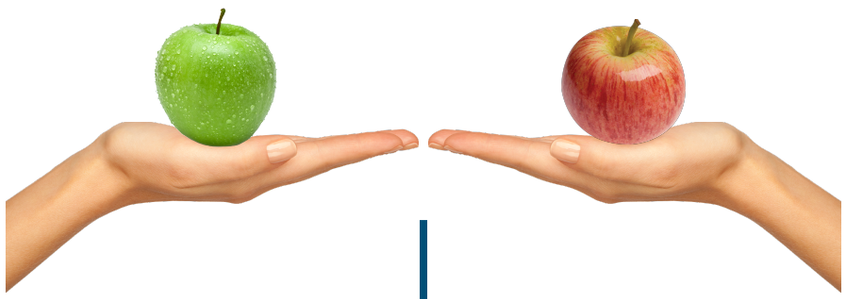
Improvement



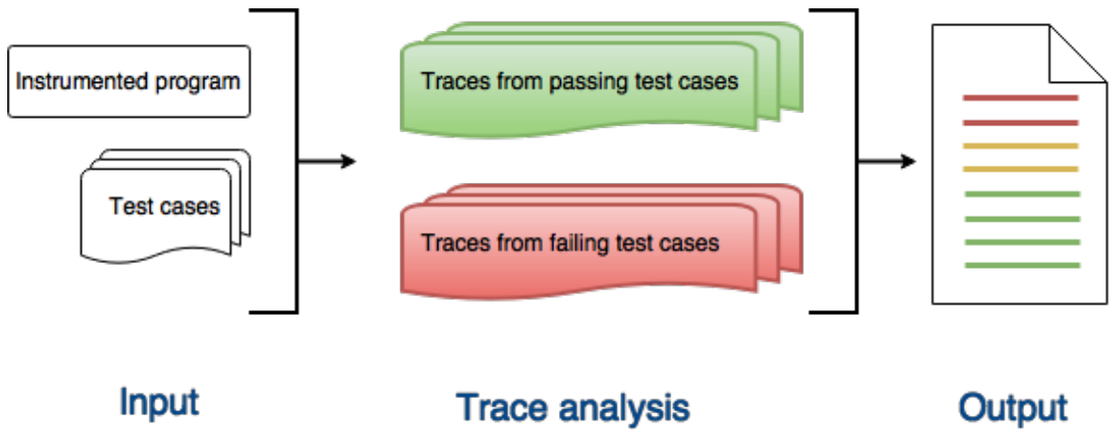
Sequenced Spectrum Analysis

Improvement

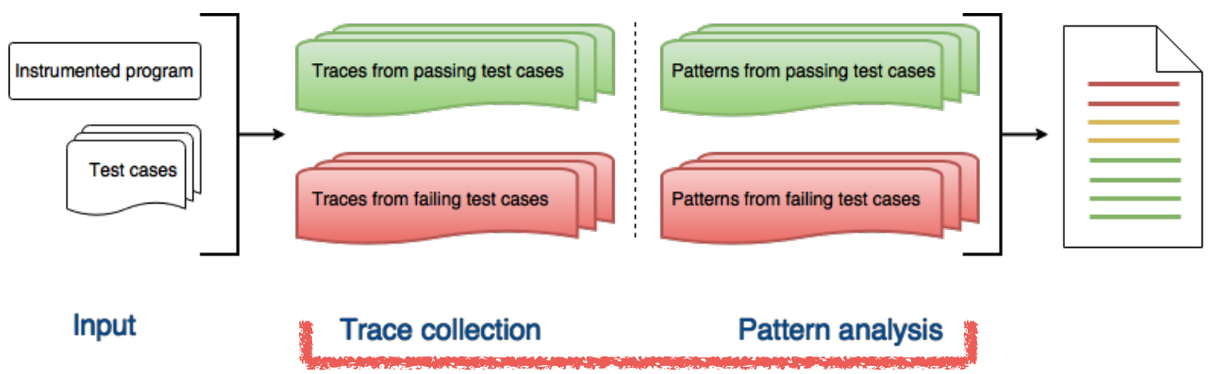
Case Study



Raw Spectrum Analysis

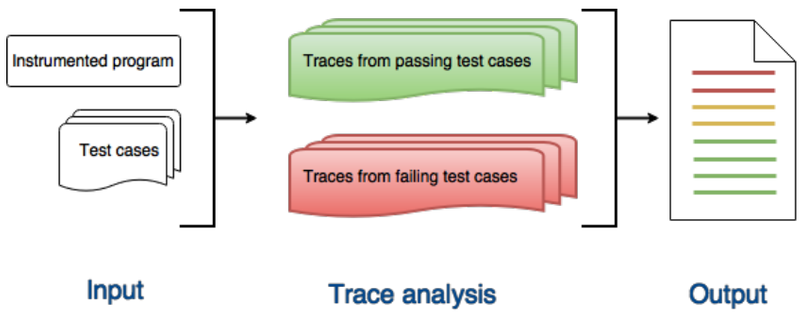


Sequenced Spectrum Analysis

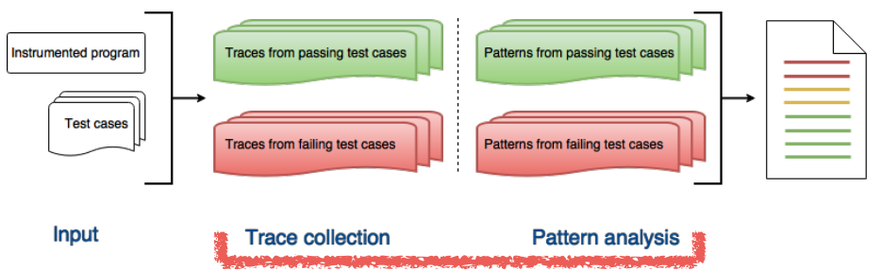


Case Study

Raw Spectrum Analysis



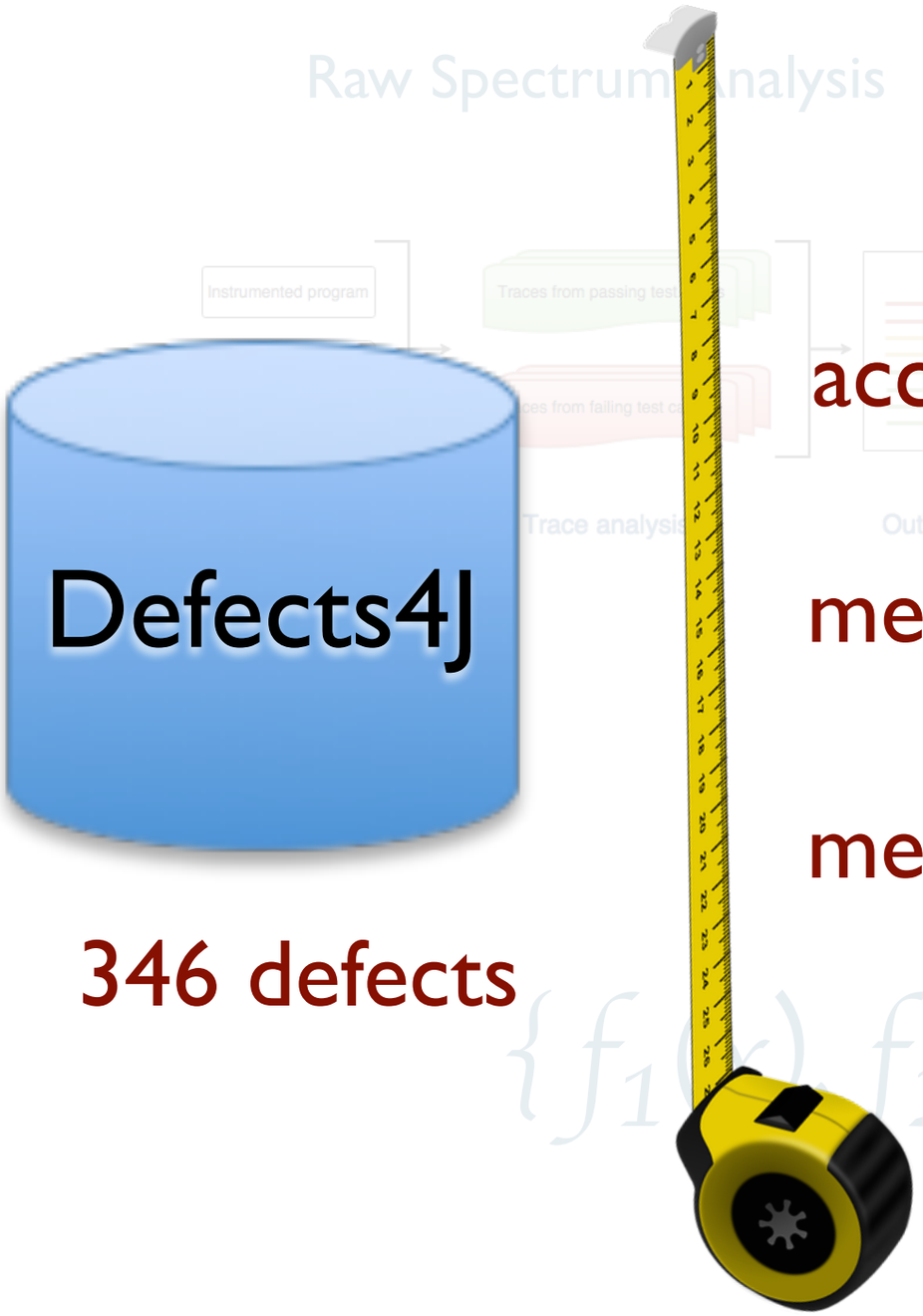
Sequenced Spectrum Analysis



+

$$\{ f_1(x), f_2(x), \ldots f_{47}(x) \}$$

Case Study



346 defects

acc@n - Higher the better

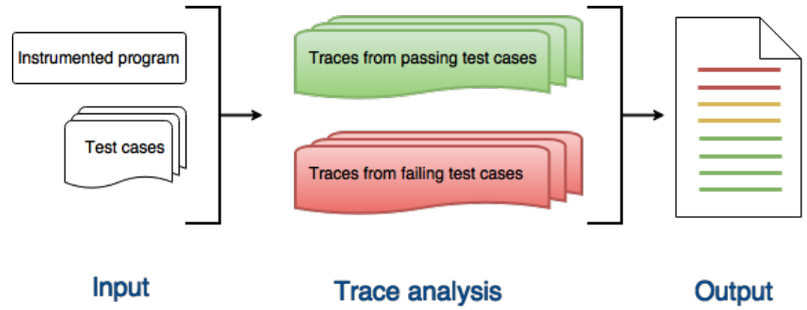
mean average precision - Higher the better

+

mean wasted effort - Lower the better

Results

Raw Spectrum Analysis



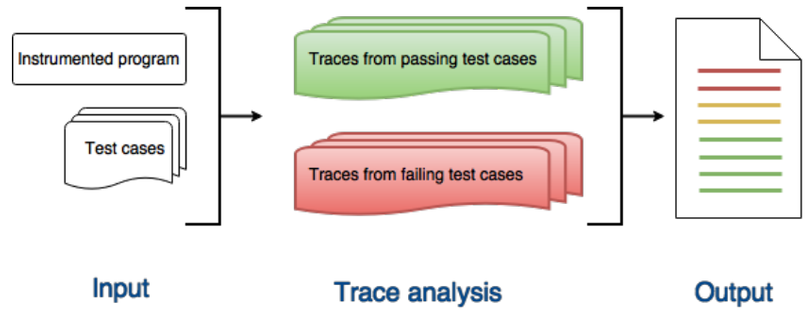
RQ1. What is the baseline performance of raw spectrum analysis?

Rank	Fault Locator	acc@1	acc@3	acc@5	MAP	MWE
1	GP13	63	120	142	0.2780349	96.73
1	Naish2	63	120	142	0.2776756	96.64
2	M2	62	118	141	0.2753030	96.32
3	Goodman	61	116	138	0.2695181	16.68
3	Ample2	64	120	140	0.2764775	101.24
4	T*	62	119	139	0.2744910	96.37
5	Zoltar	61	118	138	0.2735461	96.14

MAP = mean average precision
MWE = mean wasted effort

Results

Raw Spectrum Analysis



RQ1. What is the baseline performance of raw spectrum analysis?

Rank	Fault Locator	acc@1	acc@3	acc@5	MAP	MWE
1	GP13	63	120	142	0.2780349	96.73
1	Naish2	63	120	142	0.2776756	96.64
2		2				2
3		1				8
3	Ample2	64	120	140	0.2764775	101.24
4	T*	62	119	139	0.2744910	96.37
5	Zoltar	61	118	138	0.2735461	96.14

18% exact hit

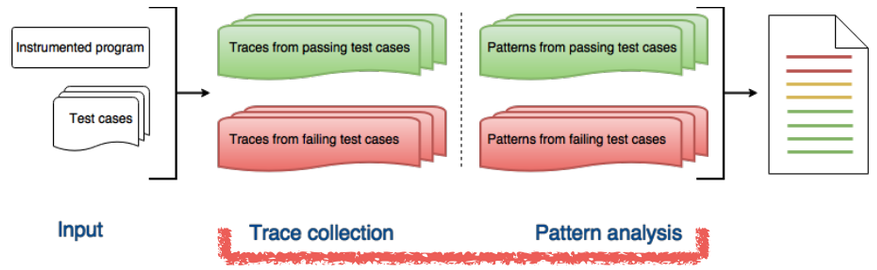
average non-faulty methods on top of the faulty method

MAP = mean average precision
MWE = mean wasted effort

Results

RQ2. How much can sequenced spectrum analysis improve upon raw spectrum analysis?

Sequenced Spectrum Analysis

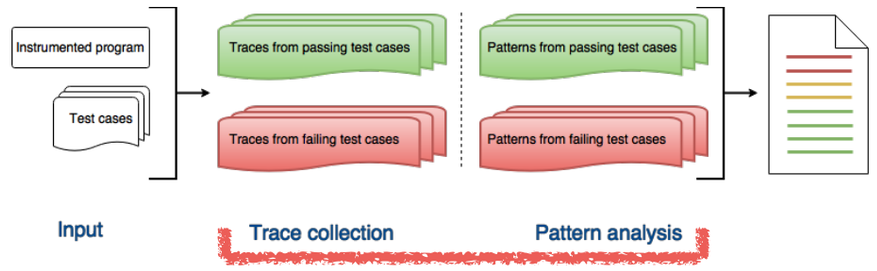


Rank	Fault Locator	acc@1	acc@3	acc@5	MAP	MWE
1	Ample2	103	159	191	0.3925699	25.88
2	Fleiss	103	157	191	0.3874628	16.01
3	T*	102	160	190	0.3936098	31.18
4	M2	102	160	189	0.3933028	31.10
5	Arithmetic Mean	103	157	189	0.3875244	26.67

MAP = mean average precision
MWE = mean wasted effort

Results

Sequenced Spectrum Analysis



RQ2. How much can sequenced spectrum analysis improve upon raw spectrum analysis?

Rank	Fault Locator	acc@1	acc@3	acc@5	MAP	MWE
1	Ample2	103	159	191	0.3925699	25.88
2	Fleiss	103	157	191	0.3874628	16.01
3	T*	102	162	189	0.3933028	31.10
4	M2	102	160	189	0.3933028	31.10
5	Arithmetic Mean	103	157	189	0.3875244	26.67

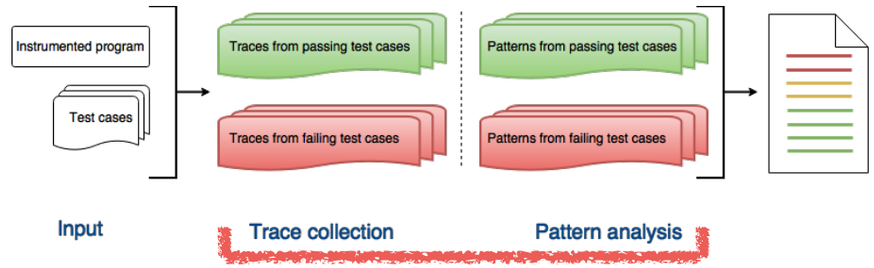
30% exact hit

average non-faulty methods on top of the faulty method

MAP = mean average precision
MWE = mean wasted effort

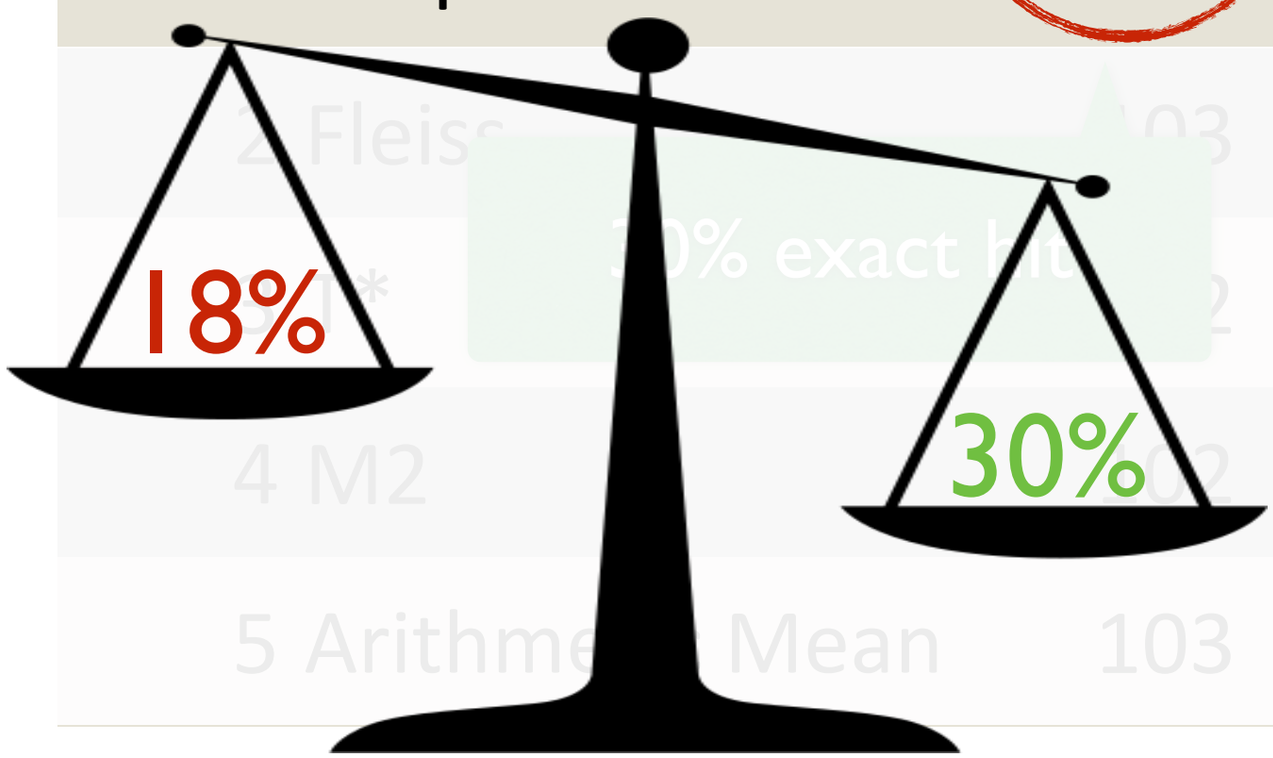
Results

Sequenced Spectrum Analysis



RQ2. How much can sequenced spectrum analysis improve upon raw spectrum analysis?

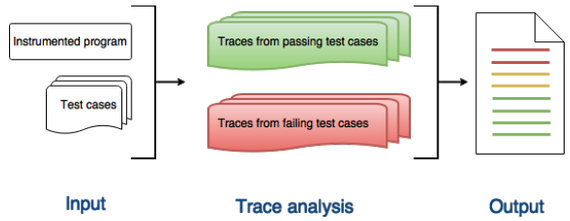
Rank	Fault Locator	acc@1	acc@3	acc@5	MAP	MWE
1	Ample2	103	159	191	0.3925699	25.88



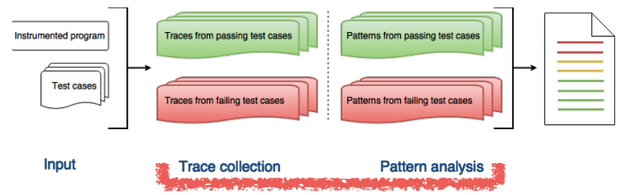
MAP = mean average precision
MWE = mean wasted effort

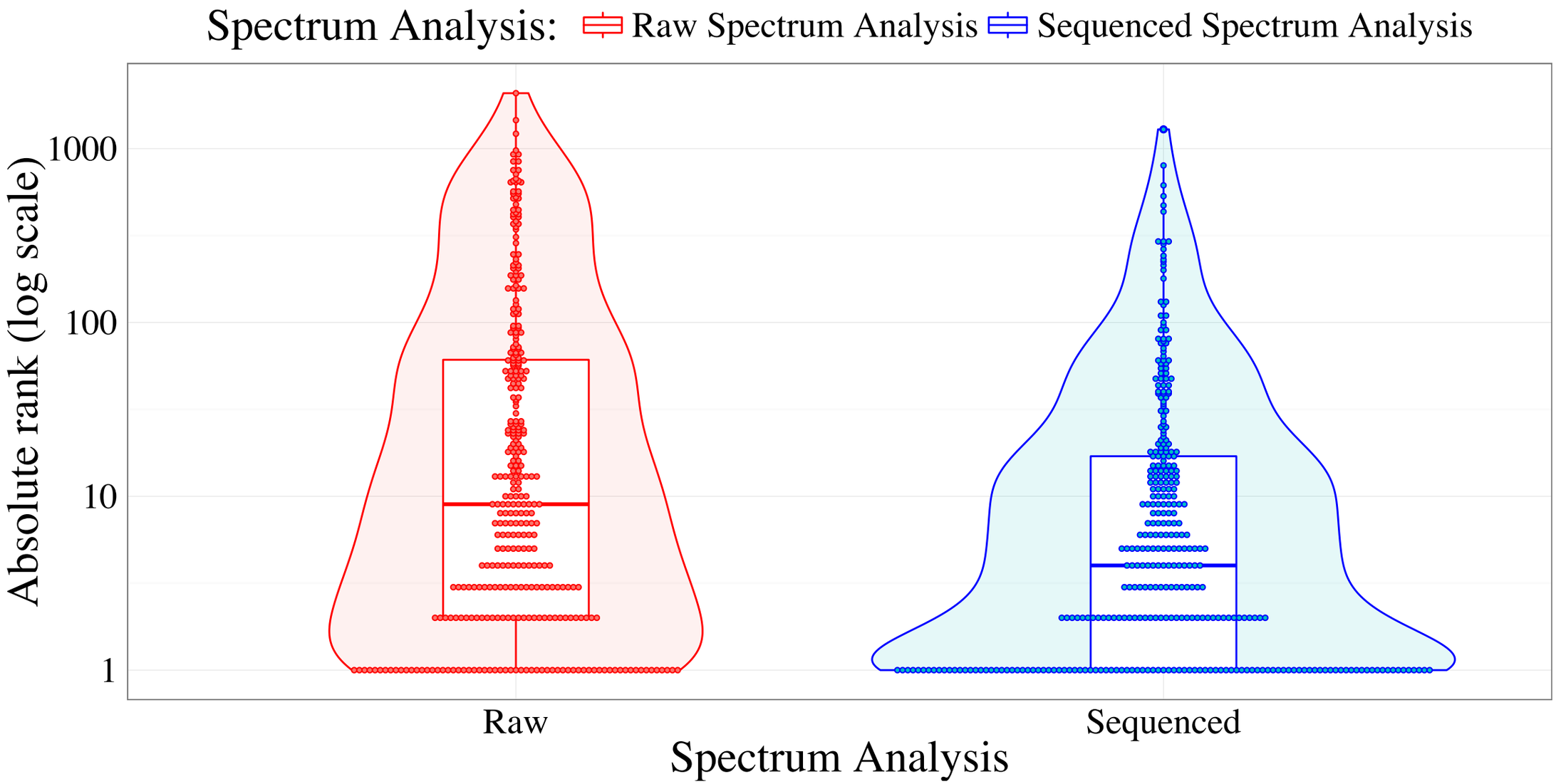
Results

Raw Spectrum Analysis

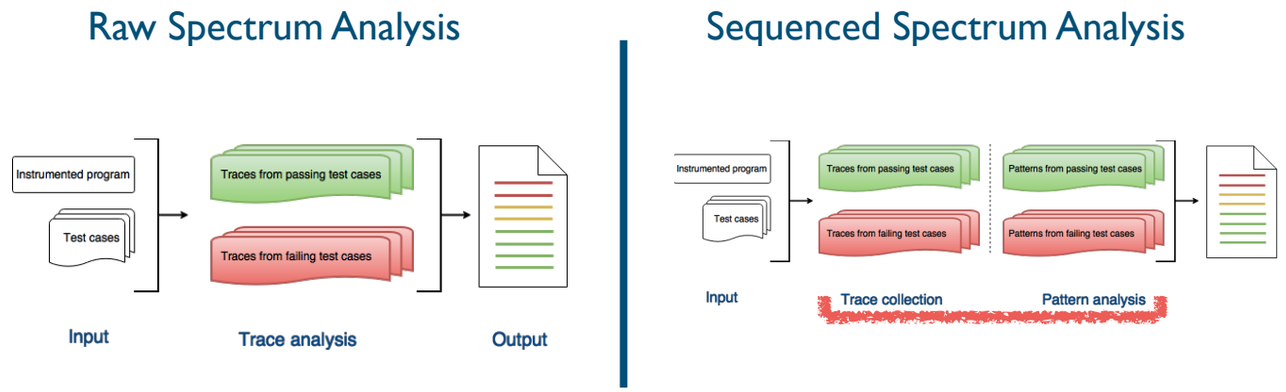


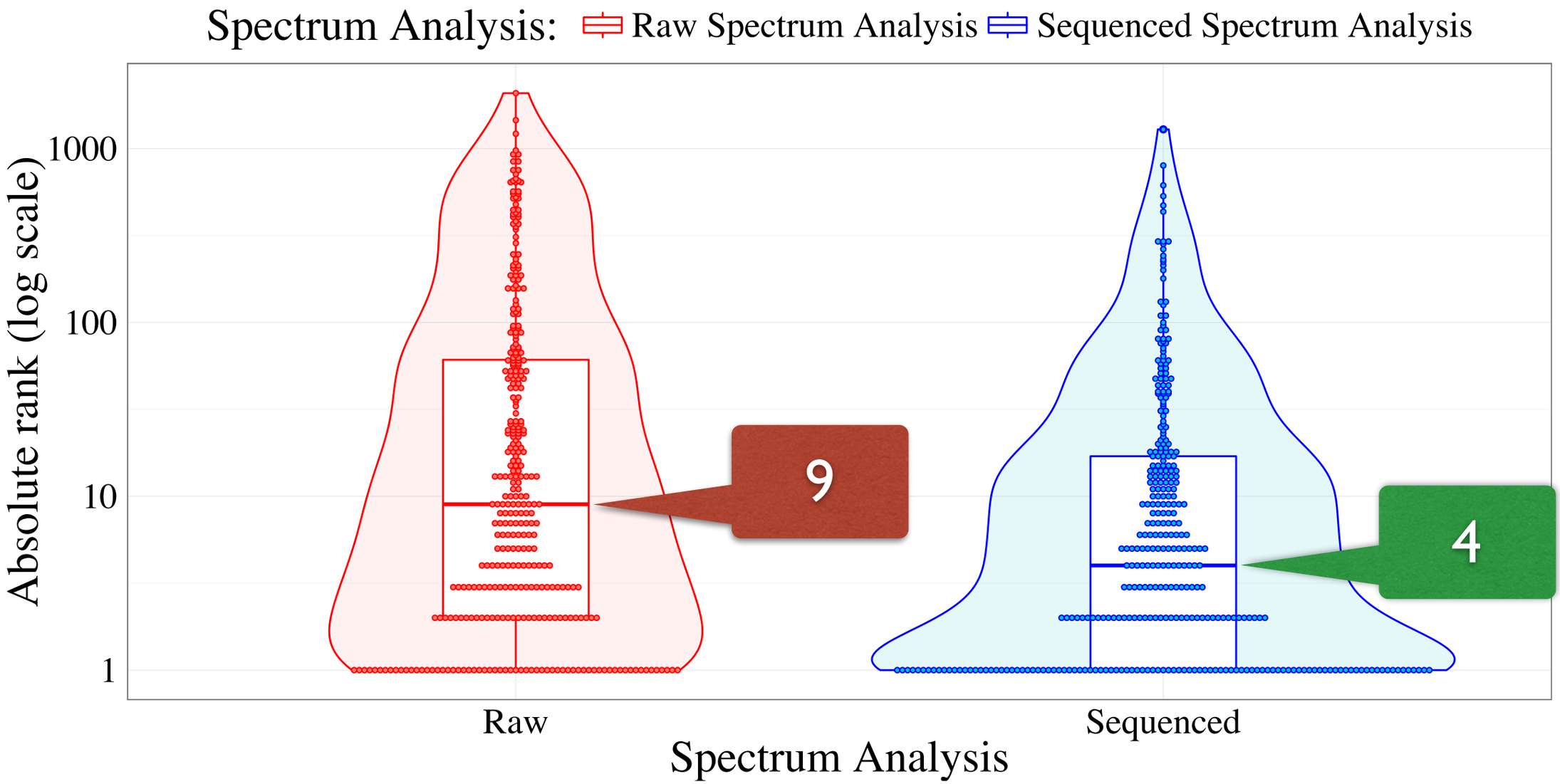
Sequenced Spectrum Analysis





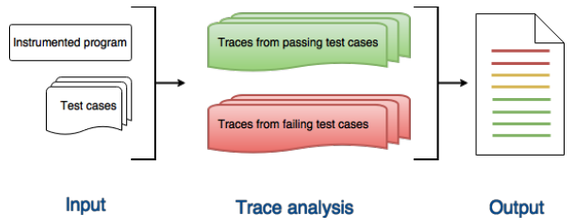
Results


Raw Spectrum Analysis
Instrumented program
Test cases
Traces from passing test cases
Traces from failing test cases
Input
Trace analysis
Output
Sequenced Spectrum Analysis
Instrumented program
Test cases
Traces from passing test cases
Traces from failing test cases
Patterns from passing test cases
Patterns from failing test cases
Input
Trace collection
Pattern analysis

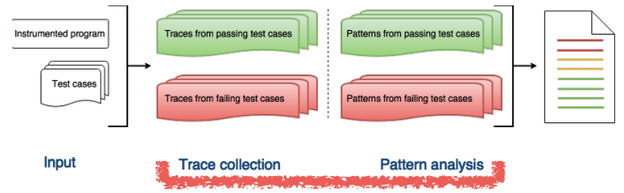

Spectrum Analysis:
Raw Spectrum Analysis
Sequenced Spectrum Analysis
Absolute rank (log scale)
1000
100
10
1
9
4
Raw
Sequenced
Spectrum Analysis

Results

Raw Spectrum Analysis



Sequenced Spectrum Analysis



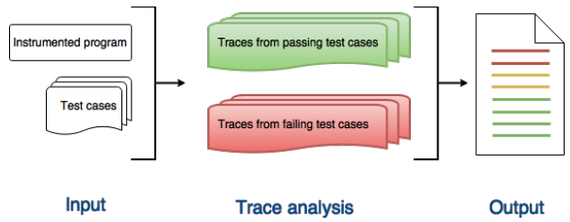
RQ3. Are there project specific differences between the rankings?

Top Fault Locators

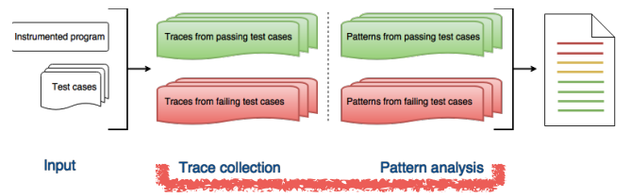
Project	Sequenced	Raw
Closure	Fleiss	GP13
Math	Goodman	Goodman
Lang	Geometric Mean	GP13
Time	Ample2	GP13
Chart	CBISqrt	Tarantula

Results

Raw Spectrum Analysis



Sequenced Spectrum Analysis



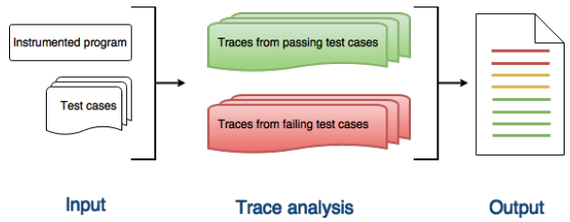
RQ3. Are there project specific differences between the rankings?

acc@1

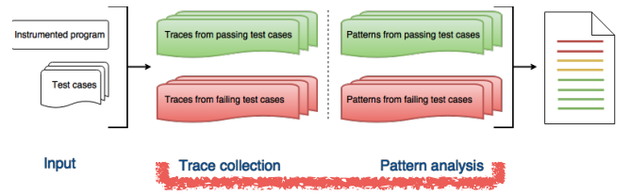
Project	Sequenced	Raw
Closure	17 / 133	7 / 133
Math	33 / 106	21 / 106
Lang	41 / 65	21 / 65
Time	6 / 27	5 / 27
Chart	9 / 26	10 / 26

Results

Raw Spectrum Analysis



Sequenced Spectrum Analysis



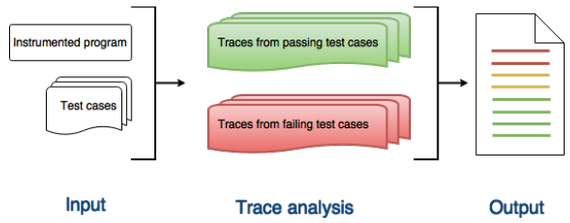
RQ3. Are there project specific differences between the rankings?

acc@1

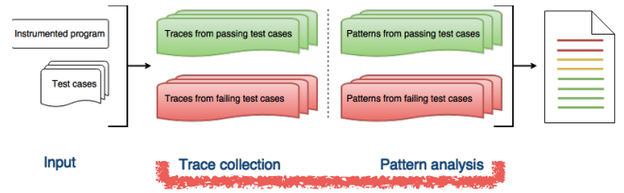
Project	Sequenced	Raw
Closure	17 / 133	7 / 133
Math	33 / 106	21 / 106
Lang	41 / 65	21 / 65
Time	6 / 27	5 / 27
Chart	9 / 26	10 / 26

Results

Raw Spectrum Analysis



Sequenced Spectrum Analysis



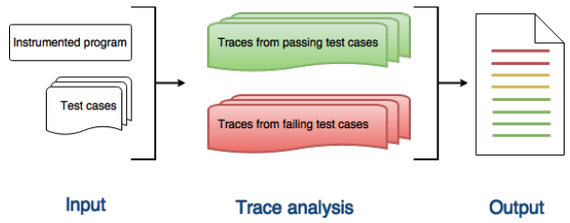
RQ3. Are there project specific differences between the rankings?

Mean Wasted Effort

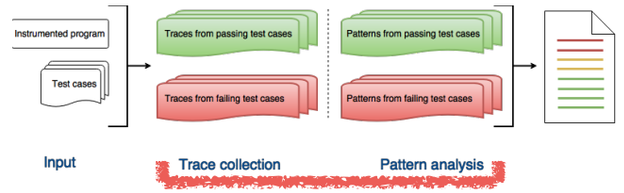
Project	Sequenced	Raw
Closure	31	223
Math	5	8
Lang	1	4
Time	16	39
Chart	13	27

Results

Raw Spectrum Analysis



Sequenced Spectrum Analysis



RQ3. Are there project specific differences between the rankings?

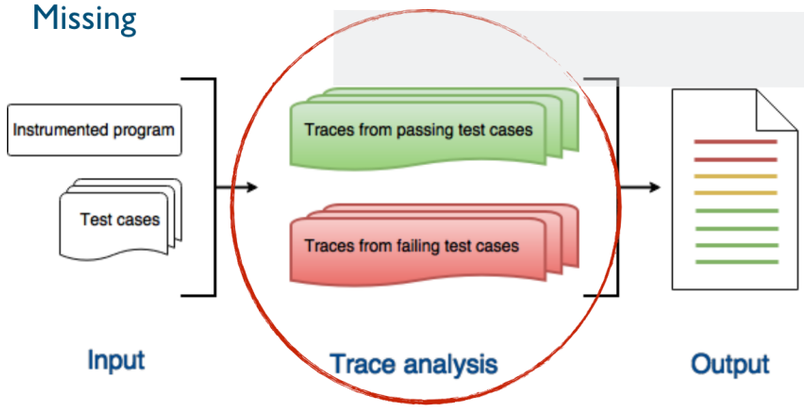
Mean Wasted Effort

Project	Sequenced	Raw
Closure	31	223
Math	5	8
Lang	1	4
Time	16	39
Chart	13	27

Summary

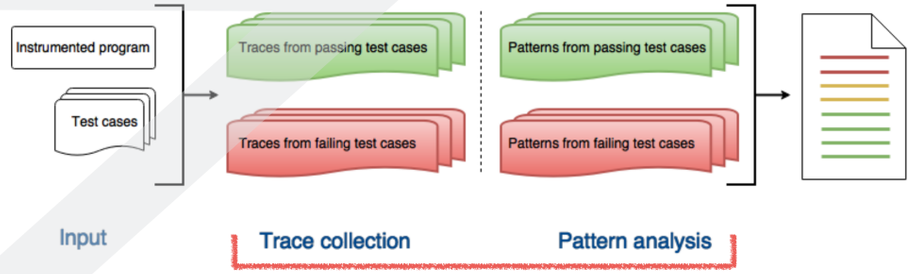
Raw Spectrum Analysis

Missing



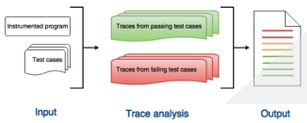
Sequenced Spectrum Analysis

Improvement



Case Study

Raw Spectrum Analysis



Sequenced Spectrum Analysis

+

$$\{ f_1(x), f_2(x), \ldots f_{47}(x) \}$$

Results