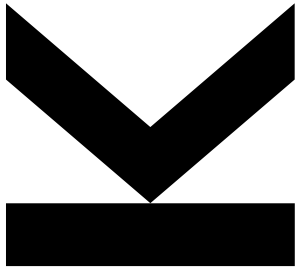


A HYBRID FEATURE LOCATION TECHNIQUE FOR RE-ENGINEERING SINGLE SYSTEMS INTO SOFTWARE PRODUCT LINES



Gabriela Karoline Michelon (Presenter)

Lukas Linsbauer

Wesley Klewerton Guez Assunção

Stefan Fischer

Alexander Egyed

FEATURE LOCATION TECHNIQUES

- **Identify** a system **artifacts** that implement a specific program functionality, a.k.a. a **feature**
- **Support** developers in various activities
 - Software **maintenance**
 - Evolution** tasks
 - Re-engineering** process

FEATURE LOCATION TECHNIQUES

Static Analysis

Depends on textual information or data flow dependencies

↓ Overestimates traces of a feature

Textual Analysis

Exploits the domain knowledge in the form of comments and identifier-names

↓ Depends on the source code naming conventions and/or the user-issue query

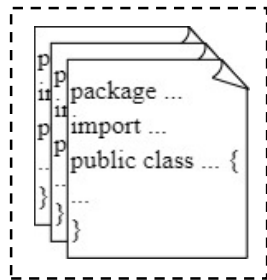
Dynamic Analysis

Monitors the features exercised

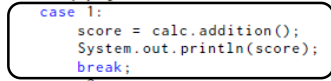
↓ Requires proper scenarios to invoke only and all traces of a specific feature

PROBLEM AND MOTIVATION

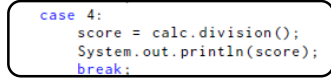
Single system



```
1 public static void main(String[] args) {
2     SimpleCalculator calc = new SimpleCalculator();
3     Scanner scan = new Scanner(System.in);
4     scan.useLocale(Locale.US);
5     int x;
6     double score;
7     calc.showMenu();
8     x = scan.nextInt();
9     switch (x) {
10        case 1:
11            score = calc.addition();
12            System.out.println(score);
13            break;
14        case 2:
15            score = calc.subtraction();
16            System.out.println(score);
17            break;
18        case 3:
19            score = calc.multiplication();
20            System.out.println(score);
21            break;
22        case 4:
23            score = calc.division();
24            System.out.println(score);
25            break;
26        default:
27            System.out.println("Wrong choice");
28            break;
29    }
30 }
31 }
```

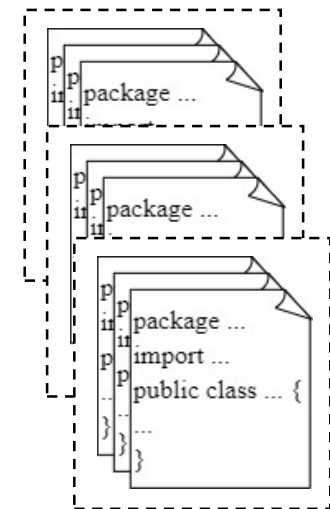


Feature Addition



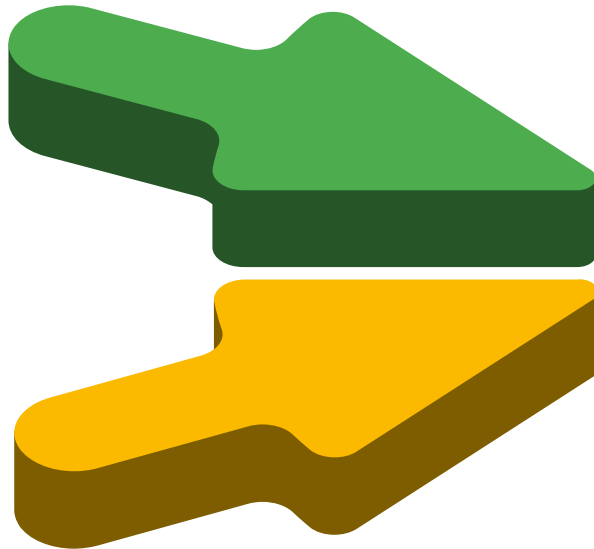
Feature Division

Variants



GOAL

- **Feature location technique (FLT)** to locate features at a **fine level** of granularity for **re-engineering a single system** into an **SPL**



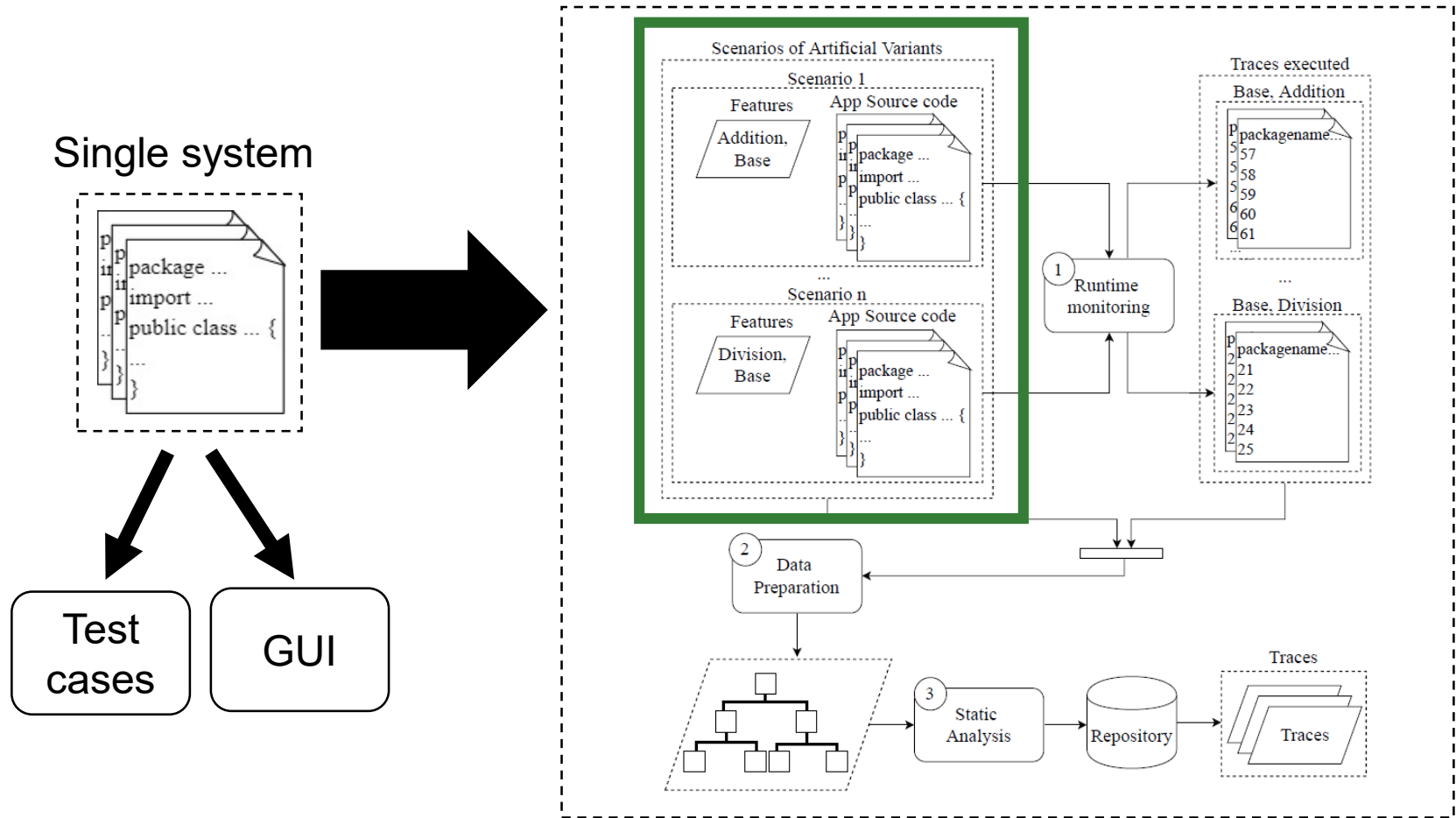
Dynamic Analysis

- Create artificial variants by exercising multiple scenarios with different features

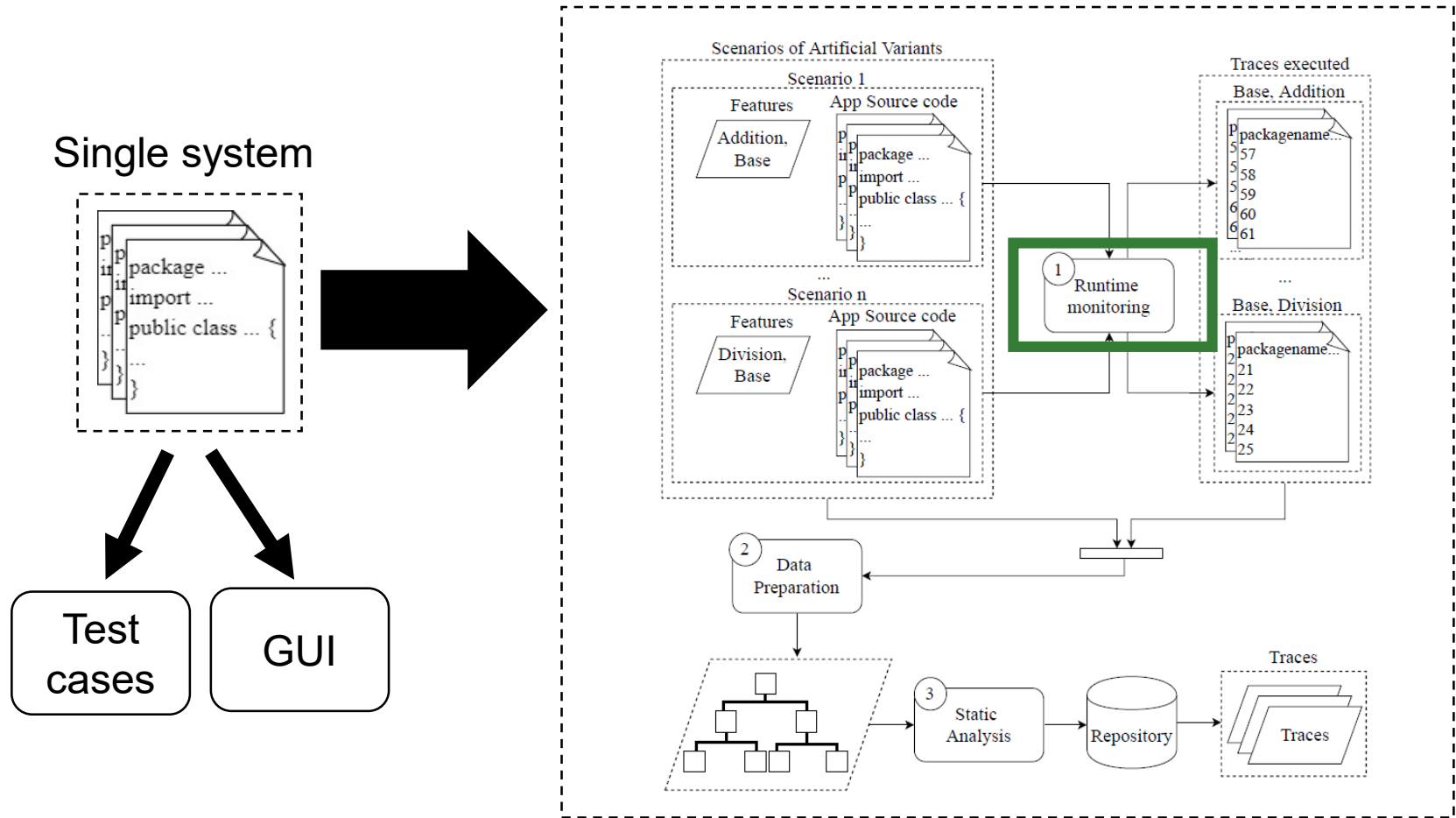
Static Analysis

- Filter the dynamic execution traces

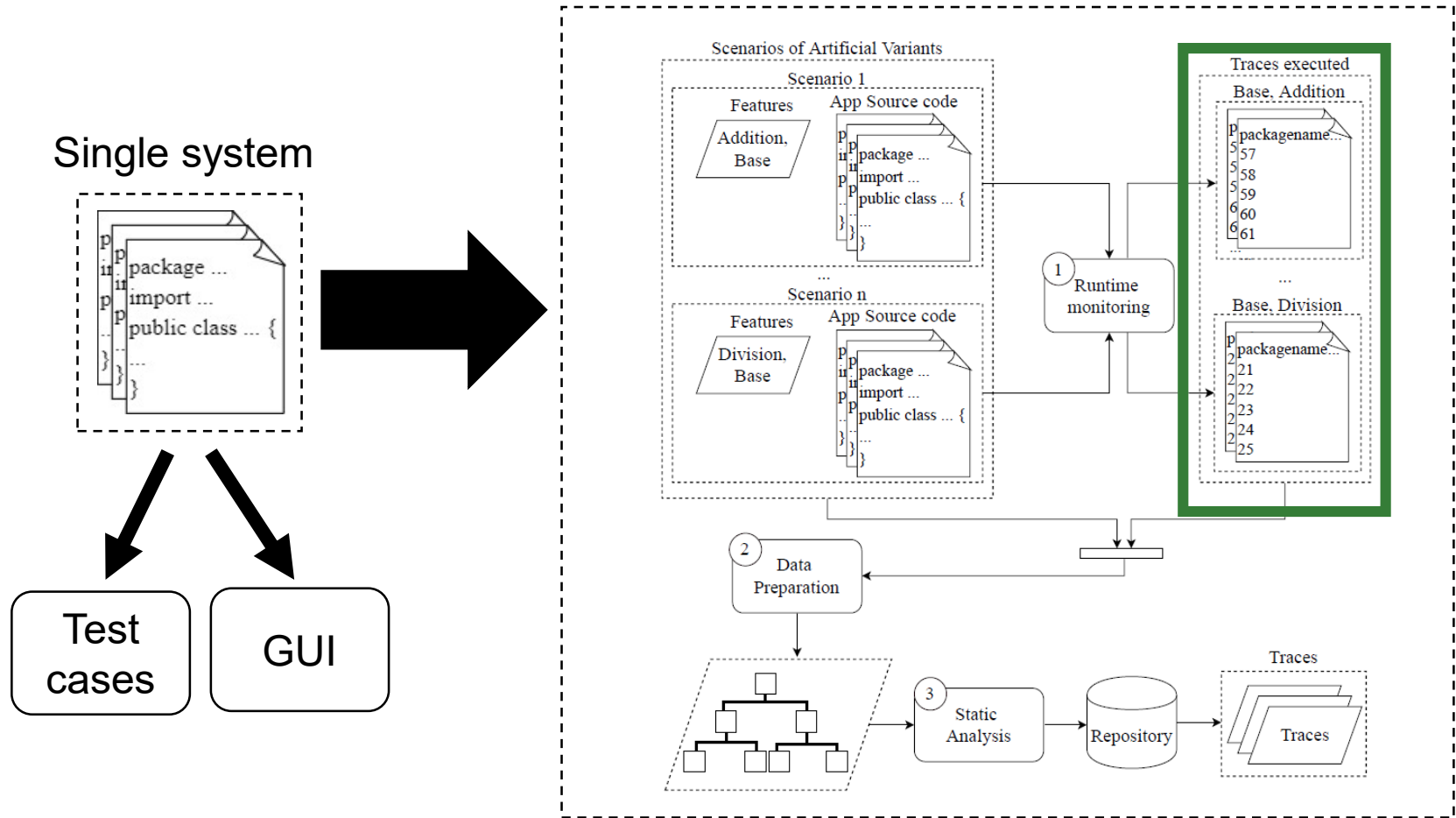
HYBRID FEATURE LOCATION TECHNIQUE



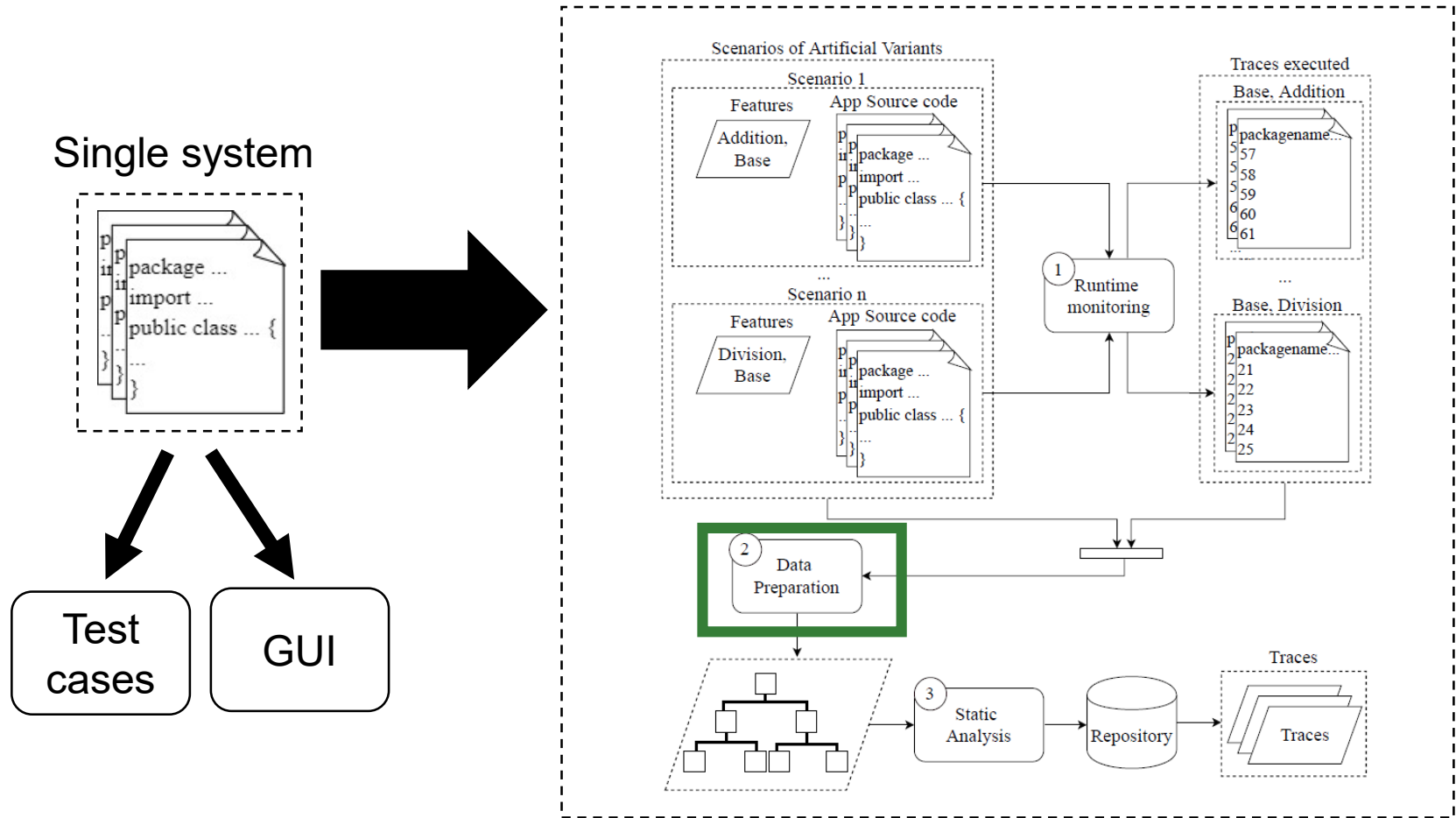
HYBRID FEATURE LOCATION TECHNIQUE



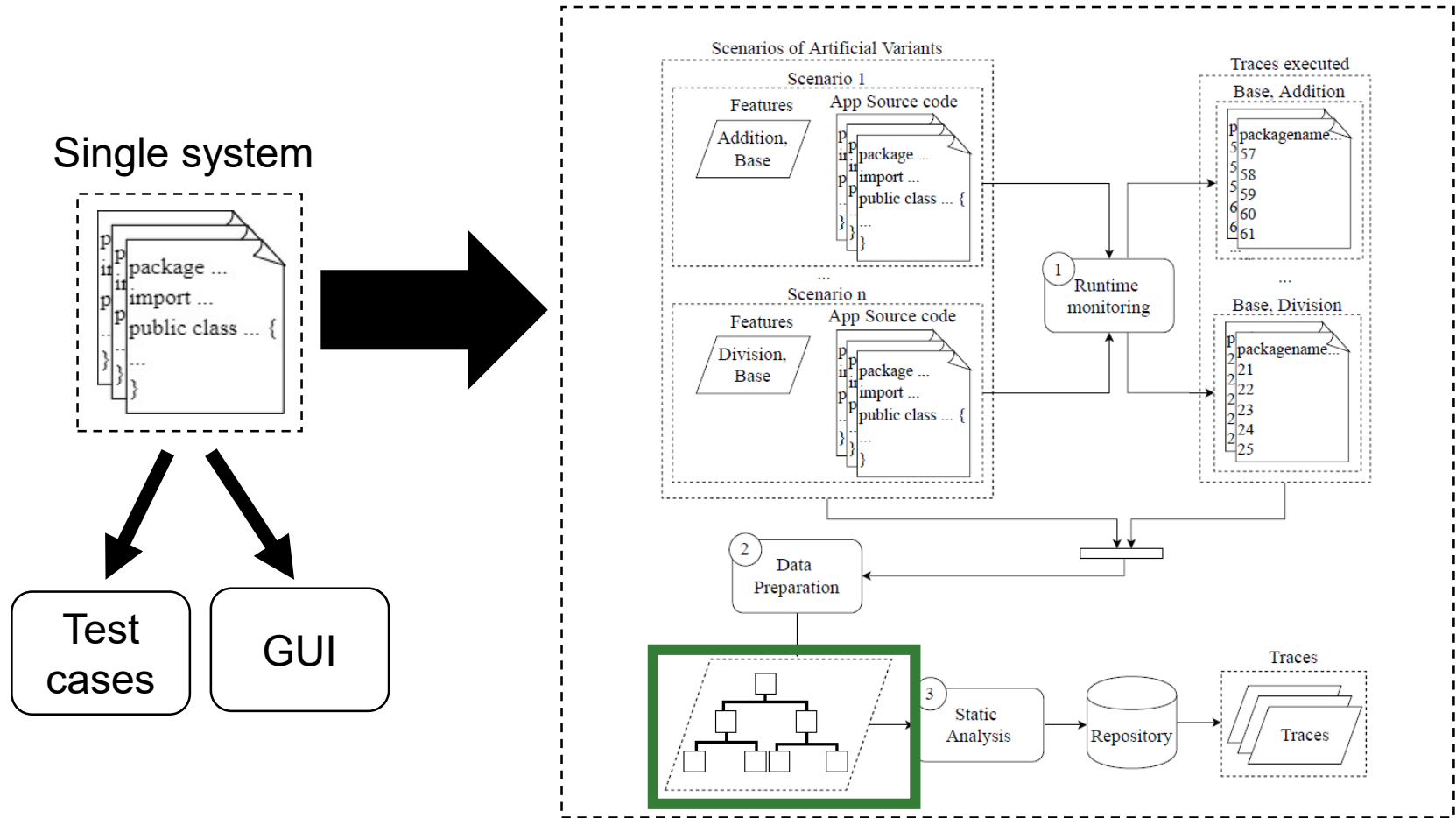
HYBRID FEATURE LOCATION TECHNIQUE



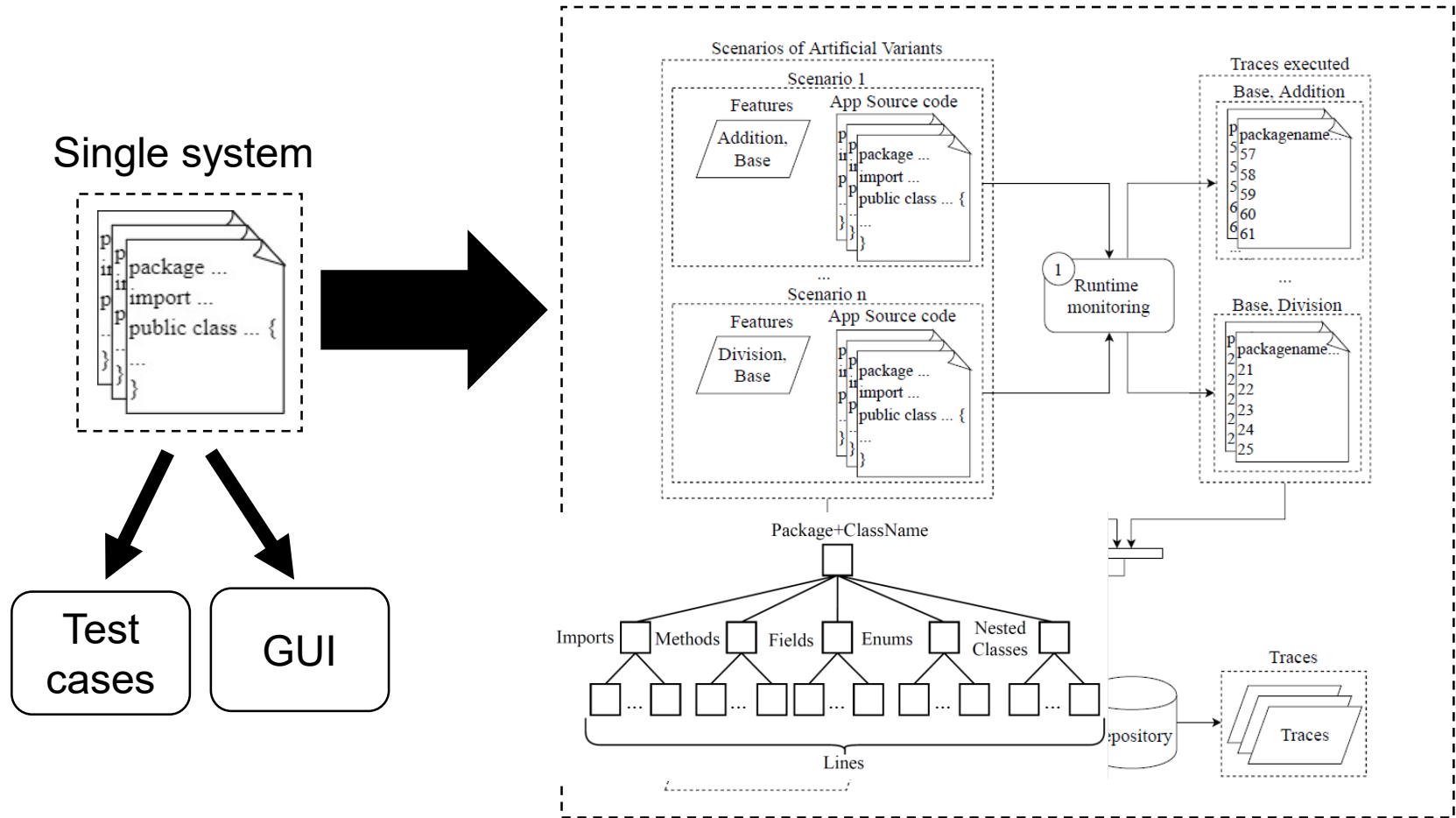
HYBRID FEATURE LOCATION TECHNIQUE



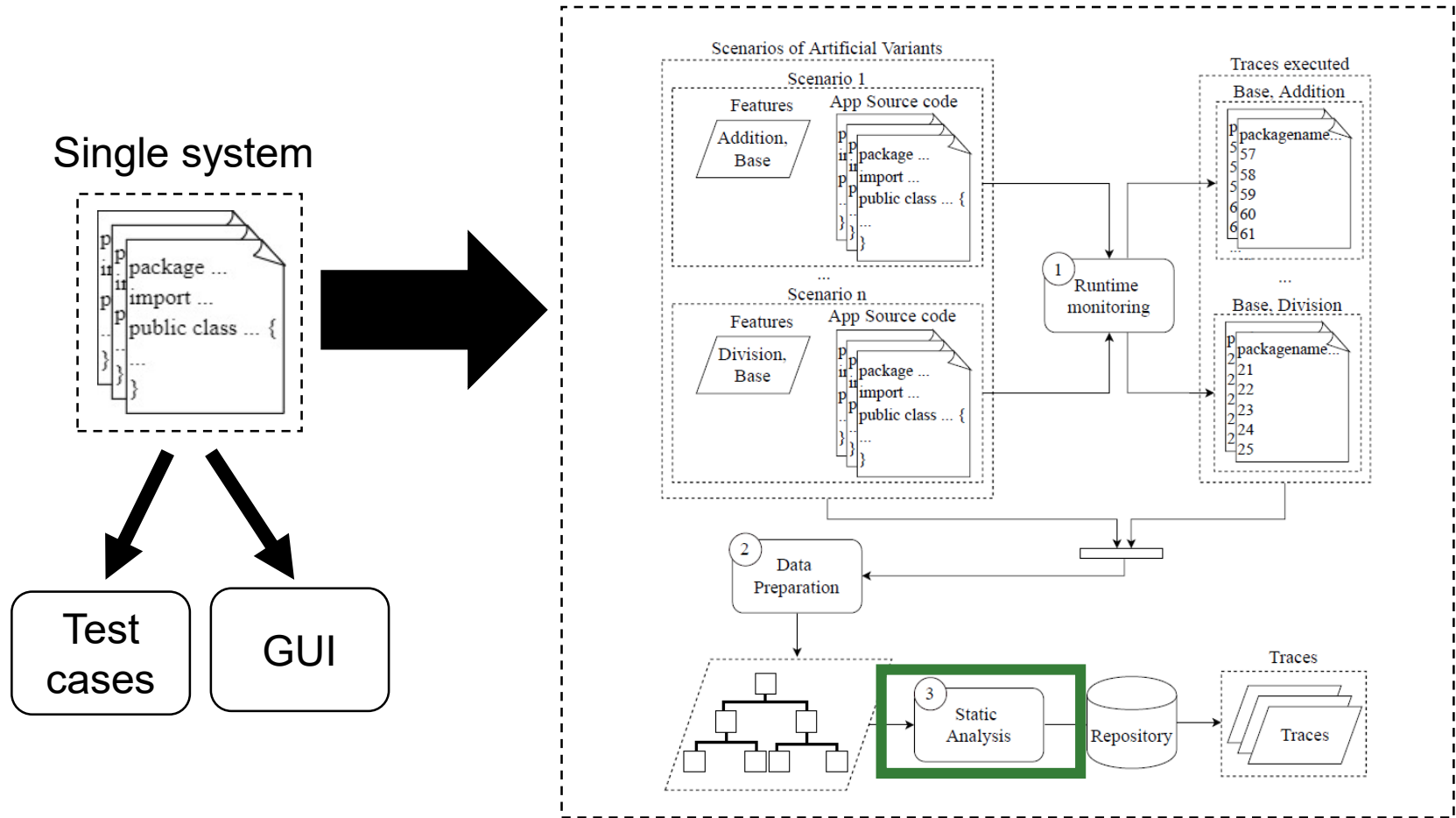
HYBRID FEATURE LOCATION TECHNIQUE



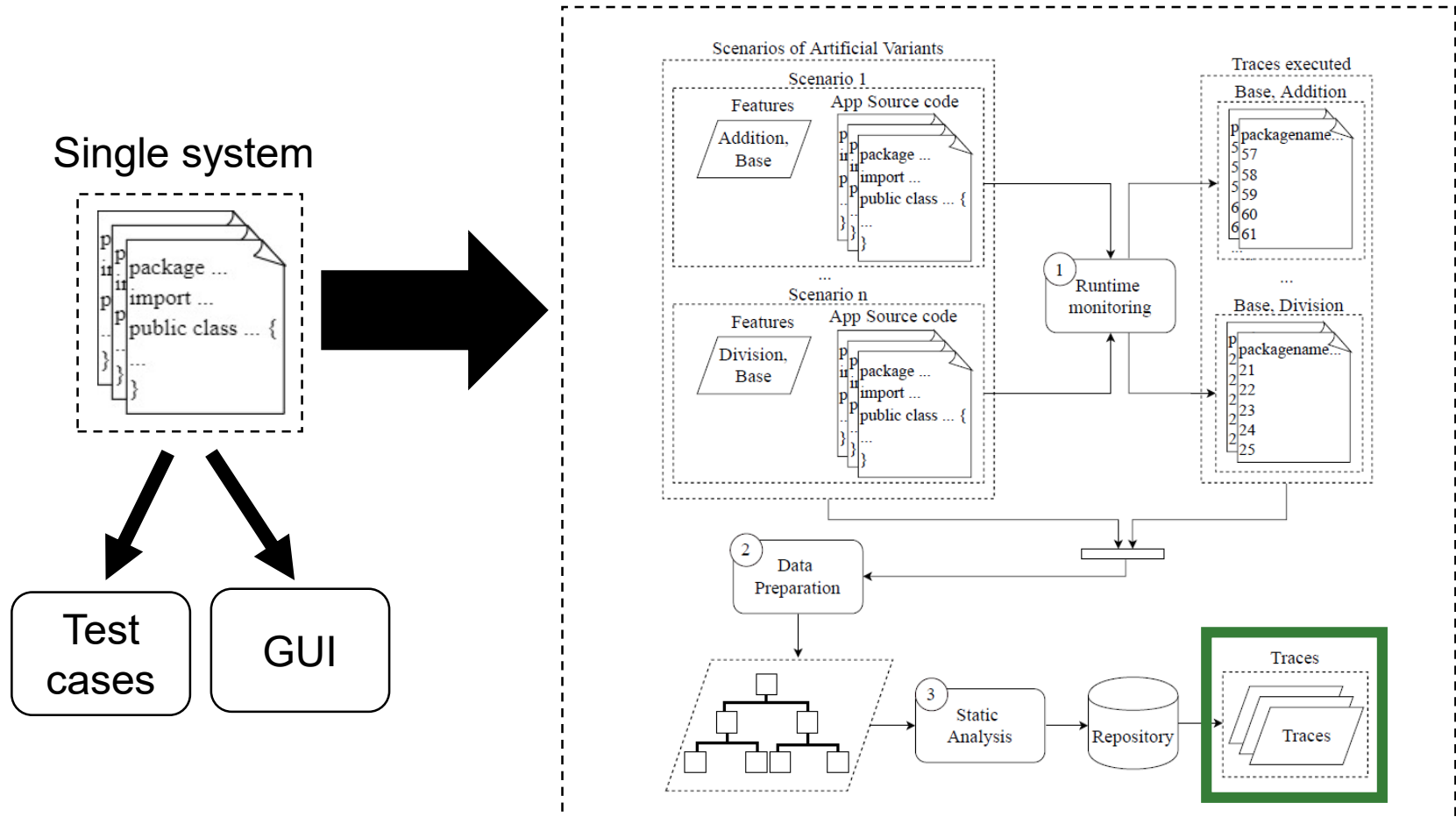
HYBRID FEATURE LOCATION TECHNIQUE



HYBRID FEATURE LOCATION TECHNIQUE



HYBRID FEATURE LOCATION TECHNIQUE



RESEARCH QUESTION

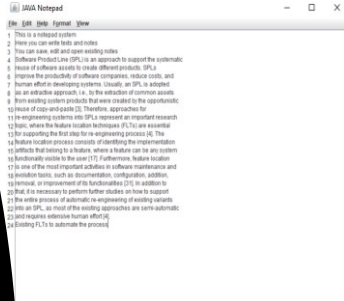


Q. How effective is
r technique for
creating features of a
ngle system?



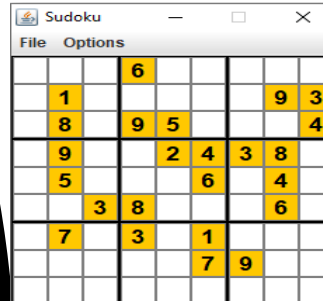
RG: Evaluate the
efficiency of our hybrid
FLT for aiding the
process of re-
engineering single
systems into SPLs

TARGET SYSTEMS



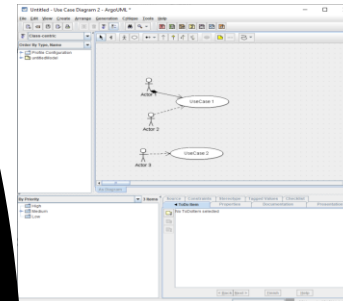
NOTEPAD

FEATURES: 20
 TEST CASES: 0
 NO GROUND
 TRUTH



SUDOKU

FEATURES: 5
 TEST CASES: 0
 NO GROUND
 TRUTH

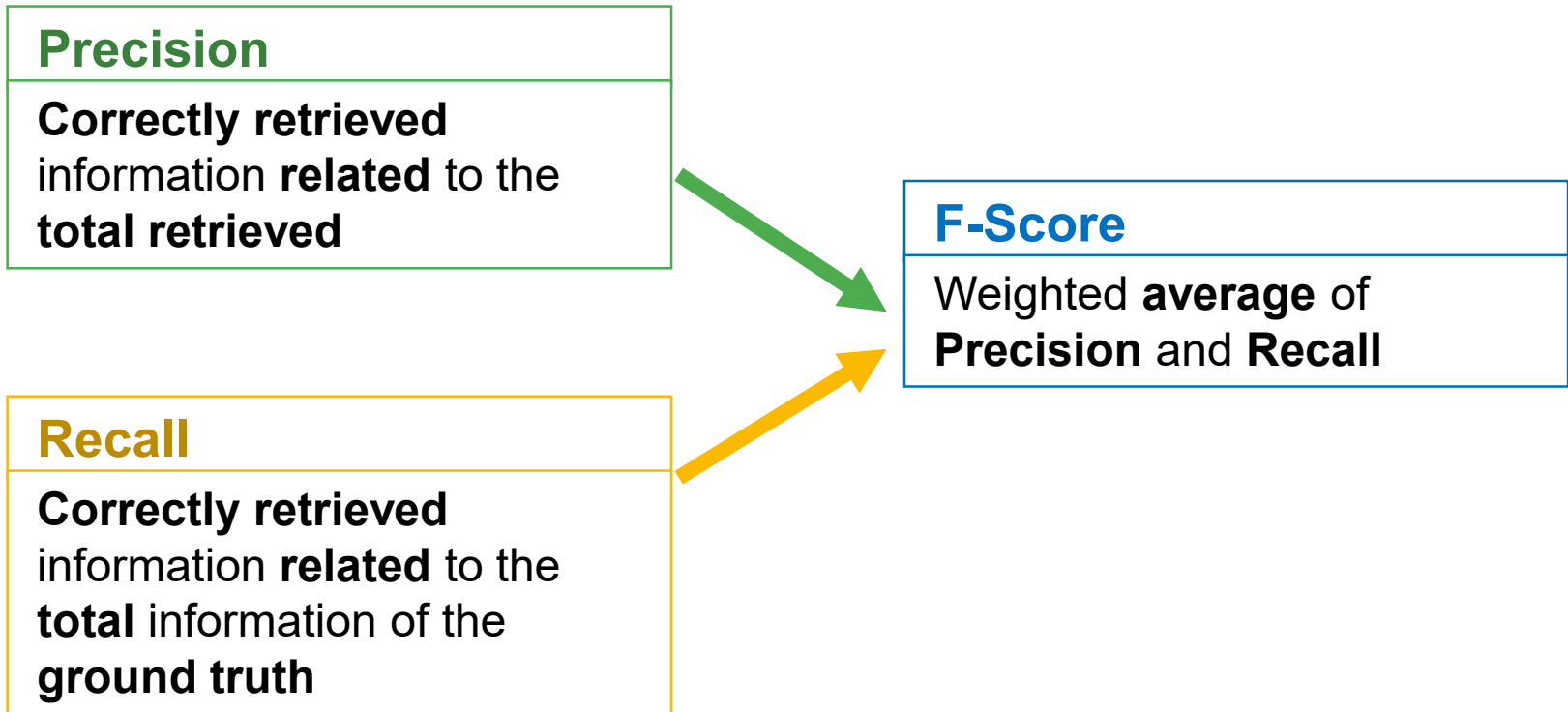


ARGOUML

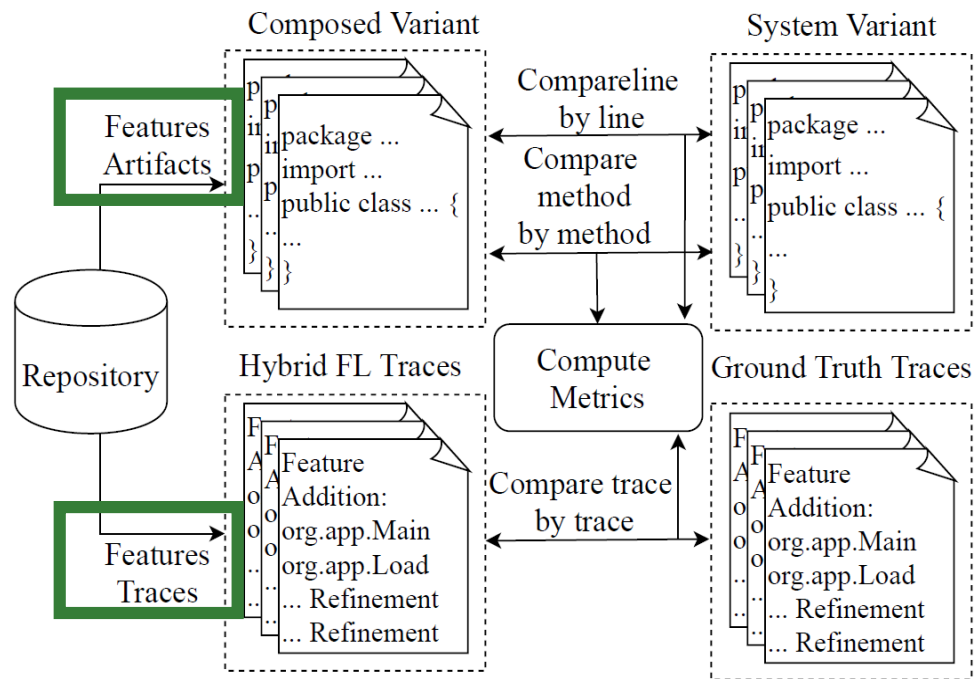
FEATURES: 8
 TEST CASES: 1198
 GROUND TRUTH
 AVAILABLE [1]

[1] Jabier Martinez, Nicolas Ordoñez, Xhevahire Tërnavá, Tewfik Ziadi, Jairo Aponte, Eduardo Figueiredo, and Marco Tulio Valente. 2018. Feature Location Benchmark with ArgoUML SPL. SPLC 2018.

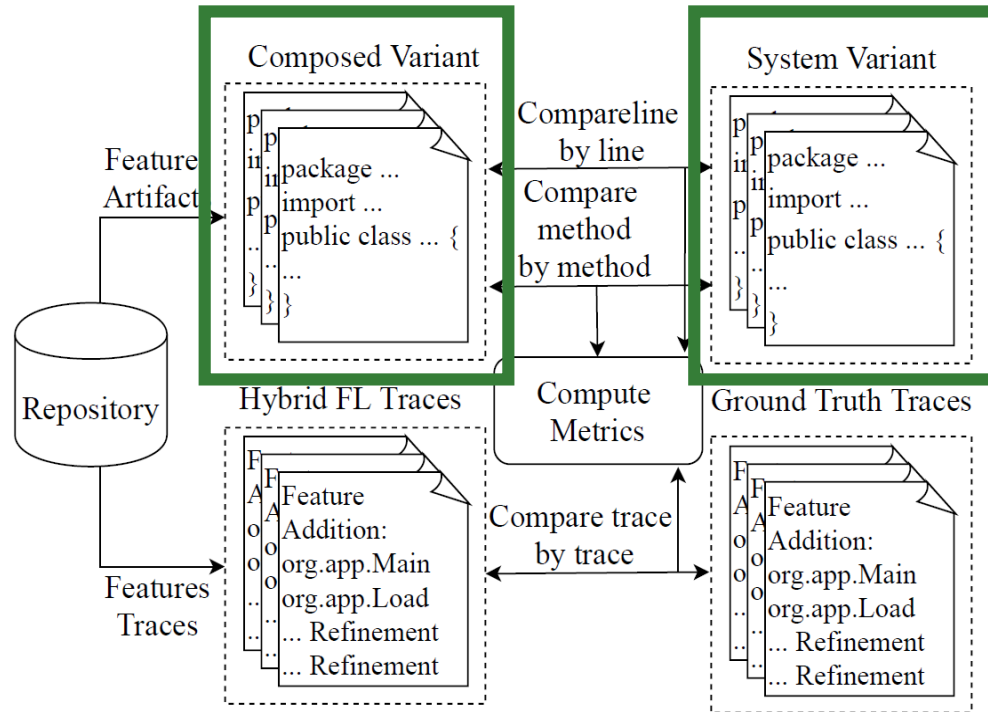
METRICS



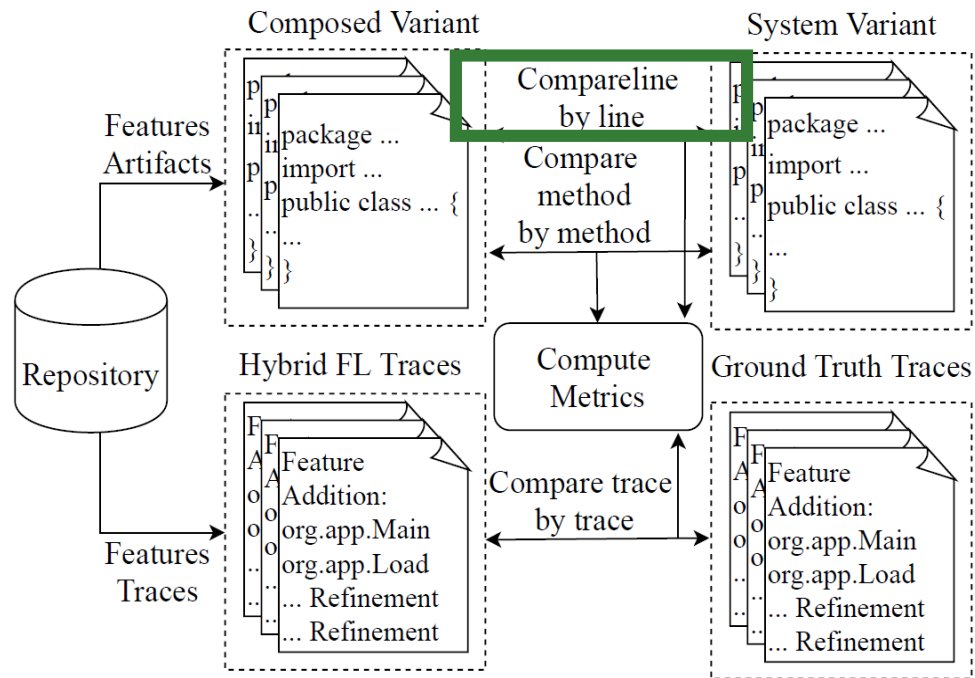
EVALUATION METHODOLOGY



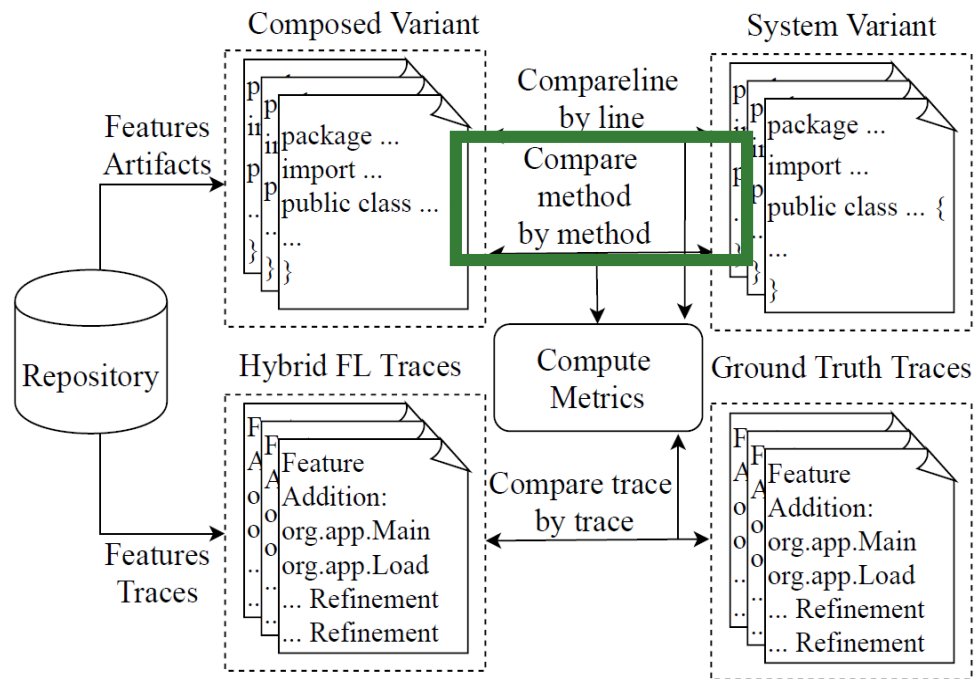
EVALUATION METHODOLOGY



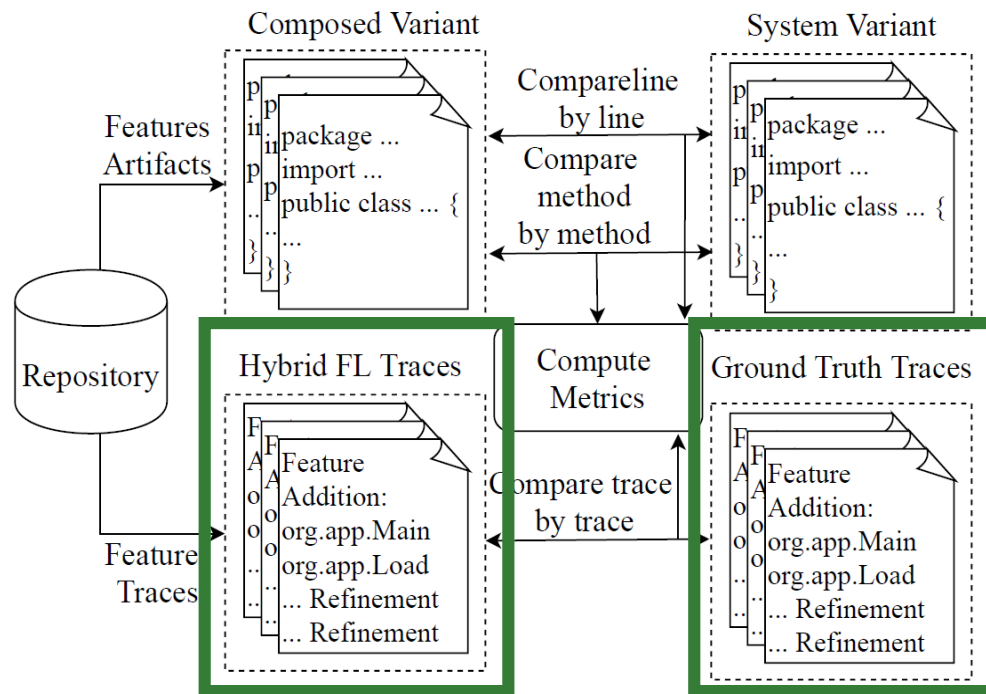
EVALUATION METHODOLOGY



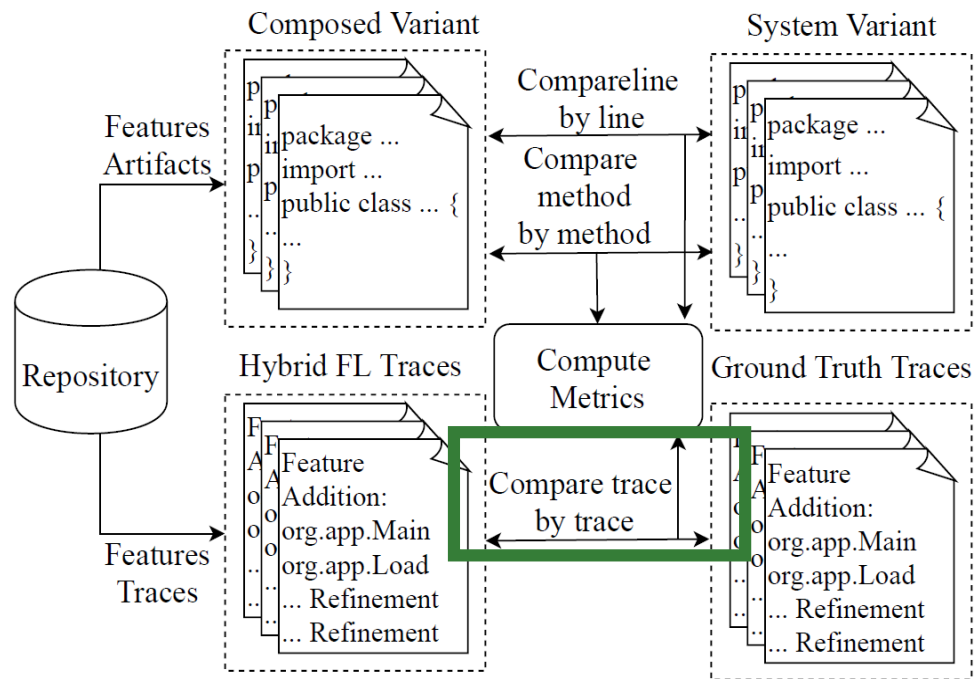
EVALUATION METHODOLOGY



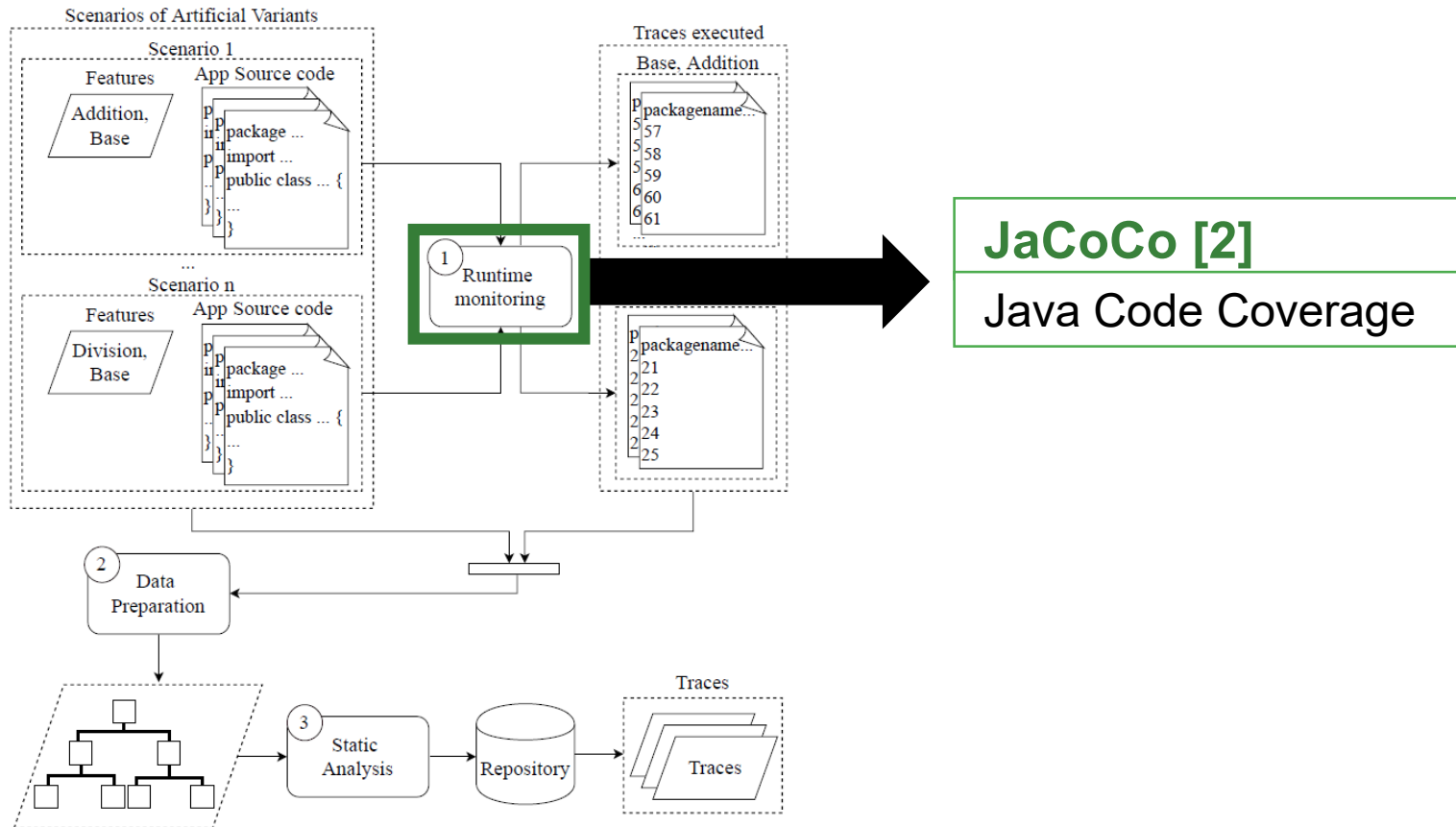
EVALUATION METHODOLOGY



EVALUATION METHODOLOGY



IMPLEMENTATION ASPECTS

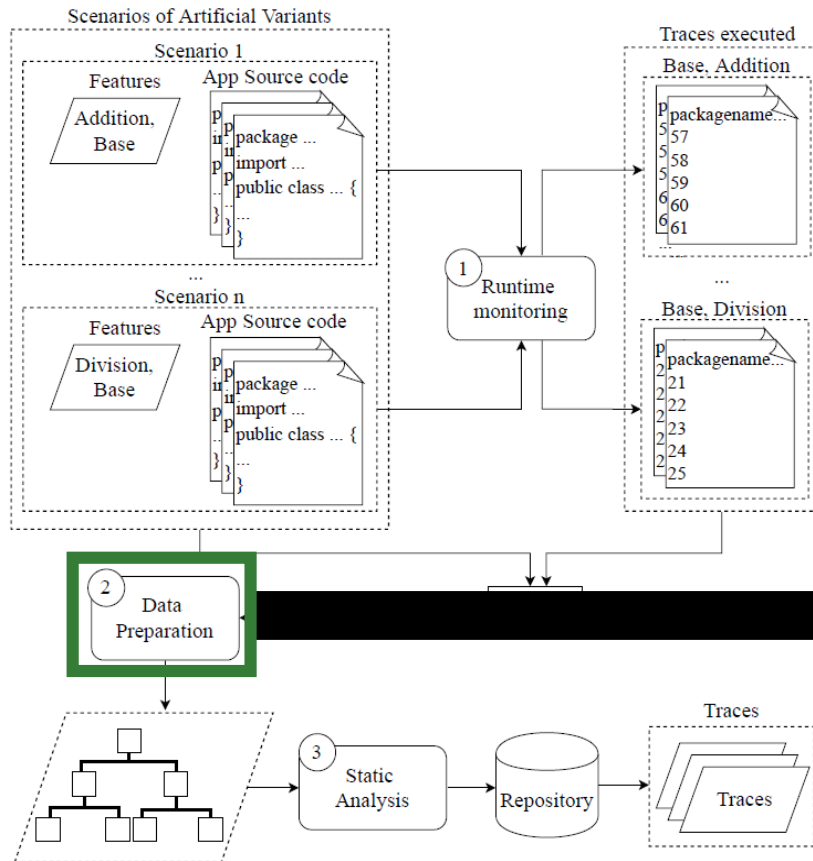


[2] <https://www.eclEmma.org/jacoco/>

[3] <https://javaparser.org/>

[4] <https://github.com/GabrielaMichelon/hybridFLT>

IMPLEMENTATION ASPECTS



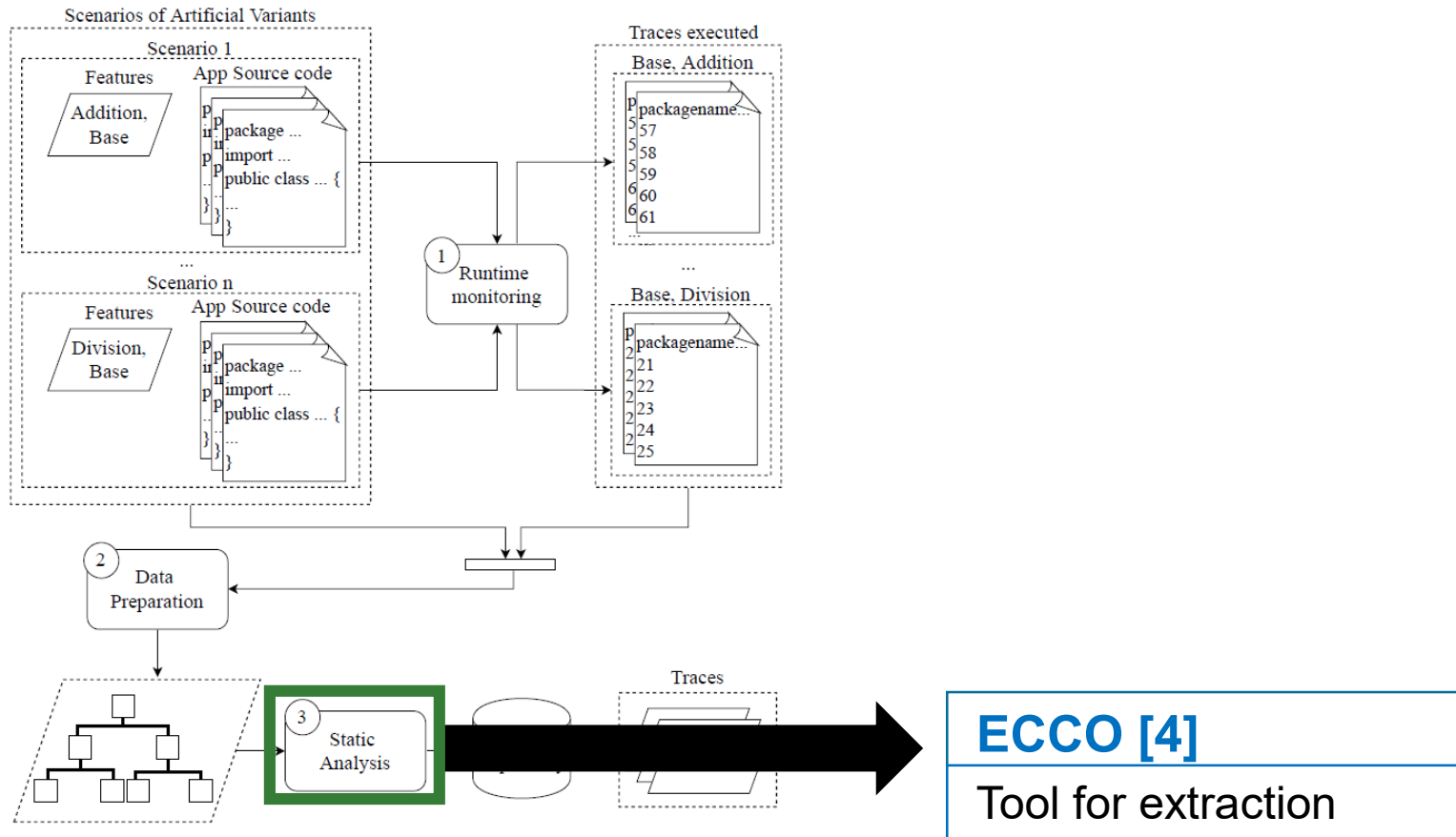
JavaParser [3]
Parser for Java

[2] <https://www.eclEmma.org/jacoco/>

[3] <https://javaparser.org/>

[4] <https://github.com/GabrielaMichelon/hybridFLT>

IMPLEMENTATION ASPECTS



[2] <https://www.eclEmma.org/jacoco/>

[3] <https://javaparser.org/>

[4] <https://github.com/GabrielaMichelon/hybridFLT>

RESULTS

Sudoku GUI

Feature	Line			Method		
	P	R	F	P	R	F
States	0.78	0.33	0.46	0.70	0.31	0.43
Solver	0.81	0.89	0.84	0.77	0.82	0.79
Generator	0.67	0.77	0.72	0.66	0.75	0.70
Undo	0.59	0.84	0.70	0.62	0.79	0.70
Extended	0.58	0.76	0.66	0.60	0.74	0.67
Average	0.68	0.72	0.70	0.67	0.68	0.68

P = Precision; R = Recall; F = F-Score.

RESULTS

Sudoku GUI

Feature	Line			Method		
	P	R	F	P	R	F
States	0.78	0.33	0.46	0.70	0.31	0.43
Solver	0.81	0.89	0.84	0.77	0.82	0.79
Generator	0.67	0.77	0.72	0.66	0.75	0.70
Undo	0.59	0.84	0.70	0.62	0.79	0.70
Extended	0.58	0.76	0.66	0.60	0.74	0.67
Average	0.68	0.72	0.70	0.67	0.68	0.68

P = Precision; R = Recall; F = F-Score.

RESULTS

Notepad GUI

Feature	Line			Method		
	P	R	F	P	R	F
About	0.78	0.77	0.77	0.60	0.36	0.45
AboutMe	0.78	0.77	0.77	0.60	0.36	0.45
Copy	0.74	0.79	0.76	0.53	0.42	0.47
Cut	0.76	0.79	0.78	0.59	0.42	0.49
ExitApp	0.81	0.74	0.77	0.67	0.42	0.51
Find	0.77	0.79	0.78	0.59	0.42	0.49
FindNext	0.78	0.78	0.78	0.59	0.40	0.48
Fonts	0.88	0.82	0.85	0.84	0.53	0.65
LineNumber	0.78	0.79	0.79	0.59	0.42	0.49
LineWrap	0.76	0.77	0.77	0.61	0.44	0.51
New	0.79	0.93	0.85	0.71	0.71	0.71
Open	0.80	0.90	0.84	0.71	0.68	0.69
Paste	0.76	0.79	0.78	0.59	0.42	0.49
Print	0.79	0.81	0.80	0.73	0.53	0.62
Redo	0.83	0.81	0.82	0.79	0.52	0.63
Save	0.78	0.94	0.86	0.72	0.75	0.73
SelectAll	0.76	0.79	0.78	0.59	0.42	0.49
TimeDate	0.76	0.79	0.78	0.59	0.42	0.49
Toolbar	0.78	0.8	0.79	0.67	0.48	0.56
Undo	0.79	0.8	0.79	0.67	0.46	0.55
Average	0.78	0.81	0.80	0.65	0.48	0.55

P = Precision; R = Recall; F = F-Score.

RESULTS

Notepad GUI

Feature	Line			Method		
	P	R	F	P	R	F
About	0.78	0.77	0.77	0.60	0.36	0.45
AboutMe	0.78	0.77	0.77	0.60	0.36	0.45
Copy	0.74	0.79	0.76	0.53	0.42	0.47
Cut	0.76	0.79	0.78	0.59	0.42	0.49
ExitApp	0.81	0.74	0.77	0.67	0.42	0.51
Find	0.77	0.79	0.78	0.59	0.42	0.49
FindNext	0.78	0.78	0.78	0.59	0.40	0.48
Fonts	0.88	0.82	0.85	0.84	0.53	0.65
LineNumber	0.78	0.79	0.79	0.59	0.42	0.49
LineWrap	0.76	0.77	0.77	0.61	0.44	0.51
New	0.79	0.93	0.85	0.71	0.71	0.71
Open	0.80	0.90	0.84	0.71	0.68	0.69
Paste	0.76	0.79	0.78	0.59	0.42	0.49
Print	0.79	0.81	0.80	0.73	0.53	0.62
Redo	0.83	0.81	0.82	0.79	0.52	0.63
Save	0.78	0.94	0.86	0.72	0.75	0.73
SelectAll	0.76	0.79	0.78	0.59	0.42	0.49
TimeDate	0.76	0.79	0.78	0.59	0.42	0.49
Toolbar	0.78	0.8	0.79	0.67	0.48	0.56
Undo	0.79	0.8	0.79	0.67	0.46	0.55
Average	0.78	0.81	0.80	0.65	0.48	0.55

P = Precision; R = Recall; F = F-Score.

RESULTS

ArgoUML

Feature	GUI								
	Line			Method			Traces		
	P	R	F	P	R	F	P	R	F
ActivityDiagram	0.81	0.30	0.44	0.81	0.21	0.34	0.05	0.24	0.08
Cognitive	0.97	0.27	0.42	0.83	0.28	0.42	0.23	0.52	0.32
CollaborationDiagram	0.81	0.29	0.42	0.79	0.23	0.35	0.04	0.19	0.06
DeploymentDiagram	0.82	0.31	0.45	0.78	0.25	0.37	0.04	0.45	0.07
Logging	0.82	0.25	0.38	0.71	0.26	0.37	0.00	0.00	0.00
SequenceDiagram	0.83	0.27	0.41	0.77	0.25	0.38	0.12	0.25	0.16
StateDiagram	0.81	0.28	0.42	0.77	0.26	0.39	0.08	0.31	0.13
UseCaseDiagram	0.82	0.29	0.42	0.73	0.26	0.39	0.08	0.47	0.13
Average	0.84	0.28	0.42	0.77	0.25	0.38	0.08	0.30	0.12

Feature	Tests								
	Line			Method			Traces		
	P	R	F	P	R	F	P	R	F
ActivityDiagram	0.99	0.05	0.09	1.00	0.03	0.07	0.24	0.07	0.11
Cognitive	0.98	0.14	0.24	0.94	0.11	0.20	0.26	0.49	0.34
CollaborationDiagram	0.99	0.05	0.09	0.98	0.04	0.08	0.29	0.14	0.19
DeploymentDiagram	0.99	0.04	0.08	0.96	0.04	0.07	0.57	0.10	0.16
Logging	0.99	0.04	0.09	0.85	0.04	0.09	0.95	0.10	0.18
SequenceDiagram	0.99	0.04	0.08	0.96	0.04	0.07	0.33	0.04	0.08
StateDiagram	0.99	0.05	0.10	0.97	0.05	0.09	0.28	0.16	0.21
UseCaseDiagram	0.99	0.05	0.09	0.88	0.05	0.09	0.62	0.20	0.30
Average	0.99	0.06	0.11	0.94	0.05	0.10	0.44	0.16	0.24

RESULTS

ArgoUML

Feature	GUI								
	Line			Method			Traces		
	P	R	F	P	R	F	P	R	F
ActivityDiagram	0.81	0.30	0.44	0.81	0.21	0.34	0.05	0.24	0.08
Cognitive	0.97	0.27	0.42	0.83	0.28	0.42	0.23	0.52	0.32
CollaborationDiagram	0.81	0.29	0.42	0.79	0.23	0.35	0.04	0.19	0.06
DeploymentDiagram	0.82	0.31	0.45	0.78	0.25	0.37	0.04	0.45	0.07
Logging	0.82	0.25	0.38	0.71	0.26	0.37	0.00	0.00	0.00
SequenceDiagram	0.83	0.27	0.41	0.77	0.25	0.38	0.12	0.25	0.16
StateDiagram	0.81	0.28	0.42	0.77	0.26	0.39	0.08	0.31	0.13
UseCaseDiagram	0.82	0.29	0.42	0.73	0.26	0.39	0.08	0.47	0.13
Average	0.84	0.28	0.42	0.77	0.25	0.38	0.08	0.30	0.12

Feature	Tests								
	Line			Method			Traces		
	P	R	F	P	R	F	P	R	F
ActivityDiagram	0.99	0.05	0.09	1.00	0.03	0.07	0.24	0.07	0.11
Cognitive	0.98	0.14	0.24	0.94	0.11	0.20	0.26	0.49	0.34
CollaborationDiagram	0.99	0.05	0.09	0.98	0.04	0.08	0.29	0.14	0.19
DeploymentDiagram	0.99	0.04	0.08	0.96	0.04	0.07	0.57	0.10	0.16
Logging	0.99	0.04	0.09	0.85	0.04	0.09	0.95	0.10	0.18
SequenceDiagram	0.99	0.04	0.08	0.96	0.04	0.07	0.33	0.04	0.08
StateDiagram	0.99	0.05	0.10	0.97	0.05	0.09	0.28	0.16	0.21
UseCaseDiagram	0.99	0.05	0.09	0.88	0.05	0.09	0.62	0.20	0.30
Average	0.99	0.06	0.11	0.94	0.05	0.10	0.44	0.16	0.24

RESULTS

ArgoUML

Feature	GUI								
	Line			Method			Traces		
	P	R	F	P	R	F	P	R	F
ActivityDiagram	0.81	0.30	0.44	0.81	0.21	0.34	0.05	0.24	0.08
Cognitive	0.97	0.27	0.42	0.83	0.28	0.42	0.23	0.52	0.32
CollaborationDiagram	0.81	0.29	0.42	0.79	0.23	0.35	0.04	0.19	0.06
DeploymentDiagram	0.82	0.31	0.45	0.78	0.25	0.37	0.04	0.45	0.07
Logging	0.82	0.25	0.38	0.71	0.26	0.37	0.00	0.00	0.00
SequenceDiagram	0.83	0.27	0.41	0.77	0.25	0.38	0.12	0.25	0.16
StateDiagram	0.81	0.28	0.42	0.77	0.26	0.39	0.08	0.31	0.13
UseCaseDiagram	0.82	0.29	0.42	0.73	0.26	0.39	0.08	0.47	0.13
Average	0.84	0.28	0.42	0.77	0.25	0.38	0.08	0.30	0.12

Feature	Tests								
	Line			Method			Traces		
	P	R	F	P	R	F	P	R	F
ActivityDiagram	0.99	0.05	0.09	1.00	0.03	0.07	0.24	0.07	0.11
Cognitive	0.98	0.14	0.24	0.94	0.11	0.20	0.26	0.49	0.34
CollaborationDiagram	0.99	0.05	0.09	0.98	0.04	0.08	0.29	0.14	0.19
DeploymentDiagram	0.99	0.04	0.08	0.96	0.04	0.07	0.57	0.10	0.16
Logging	0.99	0.04	0.09	0.85	0.04	0.09	0.95	0.10	0.18
SequenceDiagram	0.99	0.04	0.08	0.96	0.04	0.07	0.33	0.04	0.08
StateDiagram	0.99	0.05	0.10	0.97	0.05	0.09	0.28	0.16	0.21
UseCaseDiagram	0.99	0.05	0.09	0.88	0.05	0.09	0.62	0.20	0.30
Average	0.99	0.06	0.11	0.94	0.05	0.10	0.44	0.16	0.24

P = Precision; R = Recall; F = F-Score.

RESULTS

ArgoUML

Feature	GUI								
	Line			Method			Traces		
	P	R	F	P	R	F	P	R	F
ActivityDiagram	0.81	0.30	0.44	0.81	0.21	0.34	0.05	0.24	0.08
Cognitive	0.97	0.27	0.42	0.83	0.28	0.42	0.23	0.52	0.32
CollaborationDiagram	0.81	0.29	0.42	0.79	0.23	0.35	0.04	0.19	0.06
DeploymentDiagram	0.82	0.31	0.45	0.78	0.25	0.37	0.04	0.45	0.07
Logging	0.82	0.25	0.38	0.71	0.26	0.37	0.00	0.00	0.00
SequenceDiagram	0.83	0.27	0.41	0.77	0.25	0.38	0.12	0.25	0.16
StateDiagram	0.81	0.28	0.42	0.77	0.26	0.39	0.08	0.31	0.13
UseCaseDiagram	0.82	0.29	0.42	0.73	0.26	0.39	0.08	0.47	0.13
Average	0.84	0.28	0.42	0.77	0.25	0.38	0.08	0.30	0.12

Feature	Tests								
	Line			Method			Traces		
	P	R	F	P	R	F	P	R	F
ActivityDiagram	0.99	0.05	0.09	1.00	0.03	0.07	0.24	0.07	0.11
Cognitive	0.98	0.14	0.24	0.94	0.11	0.20	0.26	0.49	0.34
CollaborationDiagram	0.99	0.05	0.09	0.98	0.04	0.08	0.29	0.14	0.19
DeploymentDiagram	0.99	0.04	0.08	0.96	0.04	0.07	0.57	0.10	0.16
Logging	0.99	0.04	0.09	0.85	0.04	0.09	0.95	0.10	0.18
SequenceDiagram	0.99	0.04	0.08	0.96	0.04	0.07	0.33	0.04	0.08
StateDiagram	0.99	0.05	0.10	0.97	0.05	0.09	0.28	0.16	0.21
UseCaseDiagram	0.99	0.05	0.09	0.88	0.05	0.09	0.62	0.20	0.30
Average	0.99	0.06	0.11	0.94	0.05	0.10	0.44	0.16	0.24

P = Precision; R = Recall; F = F-Score.

RESULTS

ArgoUML

Feature	GUI								
	Line			Method			Traces		
	P	R	F	P	R	F	P	R	F
ActivityDiagram	0.81	0.30	0.44	0.81	0.21	0.34	0.05	0.24	0.08
Cognitive	0.97	0.27	0.42	0.83	0.28	0.42	0.23	0.52	0.32
CollaborationDiagram	0.81	0.29	0.42	0.79	0.23	0.35	0.04	0.19	0.06
DeploymentDiagram	0.82	0.31	0.45	0.78	0.25	0.37	0.04	0.45	0.07
Logging	0.82	0.25	0.38	0.71	0.26	0.37	0.00	0.00	0.00
SequenceDiagram	0.83	0.27	0.41	0.77	0.25	0.38	0.12	0.25	0.16
StateDiagram	0.81	0.28	0.42	0.77	0.26	0.39	0.08	0.31	0.13
UseCaseDiagram	0.82	0.29	0.42	0.73	0.26	0.39	0.08	0.47	0.13
Average	0.84	0.28	0.42	0.77	0.25	0.38	0.08	0.30	0.12

Feature	Tests								
	Line			Method			Traces		
	P	R	F	P	R	F	P	R	F
ActivityDiagram	0.99	0.05	0.09	1.00	0.03	0.07	0.24	0.07	0.11
Cognitive	0.98	0.14	0.24	0.94	0.11	0.20	0.26	0.49	0.34
CollaborationDiagram	0.99	0.05	0.09	0.98	0.04	0.08	0.29	0.14	0.19
DeploymentDiagram	0.99	0.04	0.08	0.96	0.04	0.07	0.57	0.10	0.16
Logging	0.99	0.04	0.09	0.85	0.04	0.09	0.95	0.10	0.18
SequenceDiagram	0.99	0.04	0.08	0.96	0.04	0.07	0.33	0.04	0.08
StateDiagram	0.99	0.05	0.10	0.97	0.05	0.09	0.28	0.16	0.21
UseCaseDiagram	0.99	0.05	0.09	0.88	0.05	0.09	0.62	0.20	0.30

org.argouml.uml.diagram.activity.ui.SelectionActionState getNewNode(int)

P = Precision; R = Recall; F = F-Score.

RESULTS

Statement coverage in ArgoUML

Feature	TS	GUI		Tests	
		CS	%	CS	%
ActivityDiagram	47826	11787	24.65	3528	7.37
Cognitive	53745	*	*	7924	14.74
CollaborationDiagram	46854	11153	23.80	3475	7.42
DeploymentDiagram	47005	11807	25.12	3373	7.18
Logging	47082	*	*	3395	7.21
SequenceDiagram	48613	11044	22.72	3377	6.95
StateDiagram	48327	11304	23.40	3440	7.12
UseCaseDiagram	47396	11516	24.30	2427	5.12

TS = Total Statements; CS = Covered Statements.

RESULTS

Statement coverage in ArgoUML

Feature	TS	GUI		Tests	
		CS	%	CS	%
ActivityDiagram	47826	11787	24.65	3528	7.37
Cognitive	53745	*	*	7924	14.74
CollaborationDiagram	46854	11153	23.80	3475	7.42
DeploymentDiagram	47005	11807	25.12	3373	7.18
Logging	47082	*	*	3395	7.21
SequenceDiagram	48613	11044	22.72	3377	6.95
StateDiagram	48327	11304	23.40	3440	7.12
UseCaseDiagram	47396	11516	24.30	2427	5.12

TS = Total Statements; CS = Covered Statements.

RQ: HOW EFFECTIVE IS OUR TECHNIQUE FOR LOCATING FEATURES OF A SINGLE SYSTEM?

- **Our technique** for locating features might **not be effective** in **large systems**
- We obtained **higher precision** for each **feature** of **all** the subject **systems** at the **line-level**, as well as at the **method-level**
- In **comparison** to the results of **previous work** [5, 6], **our** hybrid FLT **reached better** results on **average**

[5] Daniel Cruz, Eduardo Figueiredo, and Jabier Martinez. 2019. A Literature Review and Comparison of Three Feature Location Techniques Using ArgoUML-SPL.VaMoS 2019.

[6] Gabriela Karoline Michelon, Lukas Linsbauer, Wesley K. G. Assunção, and Alexander Egyed. 2019. Comparison-Based Feature Location in ArgoUML Variants. SPLC 2019.

SUMMARY AND TAKE HOME MESSAGE

- Our technique can help developers to **save time** and **effort** when **migrating a single system** to an **SPL**
- Our **results** can be **improved** if the **exercise of features** is **optimized**
- Importance of having **tests cases** available
- **Need for**
 - a more **efficient FLT fine-grained** for a **single system**
 - exploiting **hybrid FLT**s
 - benchmarks** with **fine-level** of **granularity**
 - evaluation of FLT**s with the most **efficiently** and **common metrics**

FUTURE WORK

- **Analyze the effort for executing scenarios**
- **Evaluate** our hybrid FLT of **how much time** and **effort** it would be **still required** by developers to **complete** the **product** after applying the technique

DATA AVAILABLE

■ Implementation

- runtime adapter** added to the **ECCO tool** to perform the Hybrid Feature Location

■ Dataset

- artificial variants** and their **configurations**
- ECCO repository** containing the **traces**
- ground truth** and **composed variants**
- metrics results**
- videos** recorded from **exercising features**



A HYBRID FEATURE LOCATION TECHNIQUE FOR RE-ENGINEERING SINGLE SYSTEMS INTO SOFTWARE PRODUCT LINES

THANK YOU!

Contact gabriela.michelon@jku.at

Dataset

