





How I Met Your Implemented Variability: Identification in Object-Oriented Systems with symfinder

Johann Mortara – Philippe Collet

Université Côte d'Azur, CNRS, I3S, France

Tutorial at SPLC '21 September 6, 2021

Thanks to SPLC'21 sponsors!



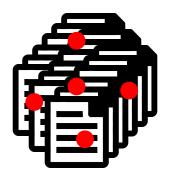






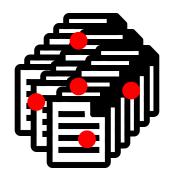


What this is all about



Variability-rich system in a <u>single</u> code base

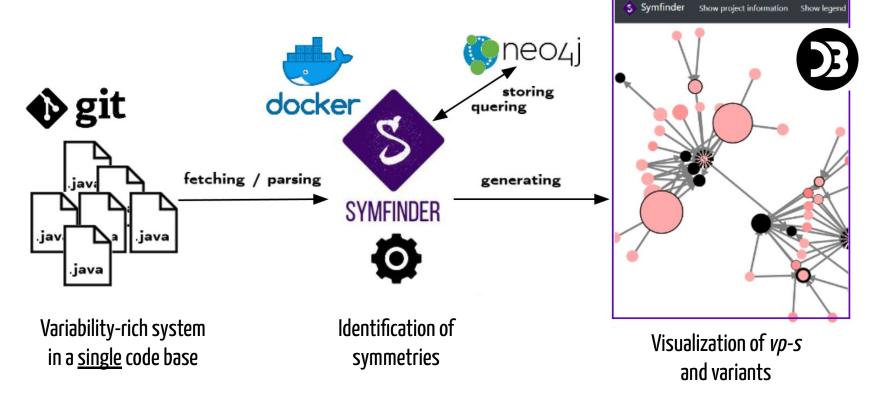
What this is all about



Variability-rich system in a <u>single</u> code base

How to identify these variability implementations?

What this is all about



1) Motivation and introduction to symfinder (12.00 – 12.40)

- **2)** symfinder: first contact (12.40 13.30)
- **3)** Break (13.30 14.00)
- **4)** Guided use of symfinder (14.00 15.00)
- 5) Wrapping up and exchange time (15.00 15.30)

Table of contents

Motivation and introduction to symfinder

Introduction

Many known software systems are highly-variable



16.000 options managed in 25M LoC [Acher2018]

#ifdef



24.000 different platforms in 2015 [Open2015]

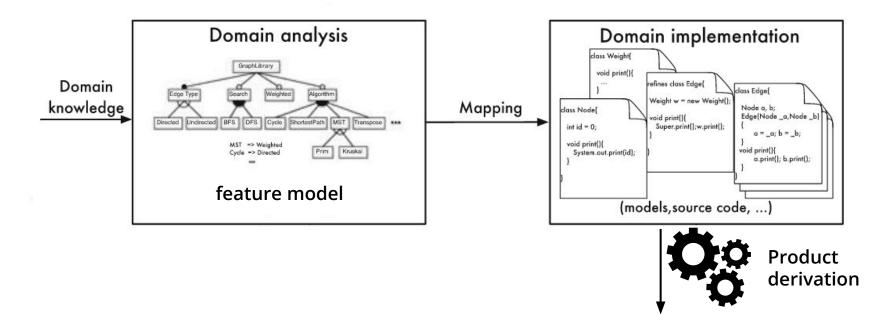


2.000+ options generating variants for platforms, security levels... [Acher2018]

Object-orientation

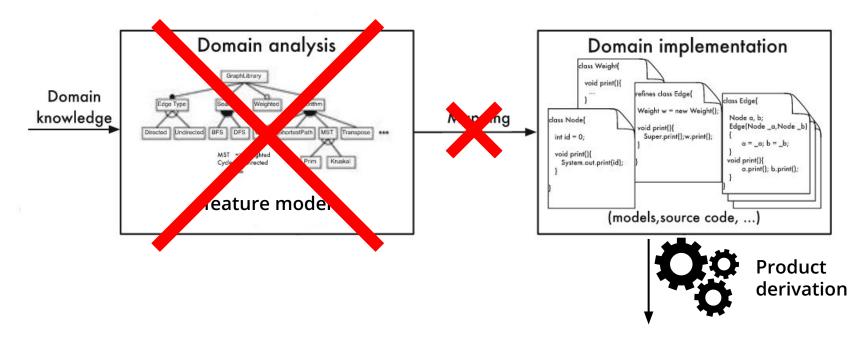
Object-orientation

Software Product Lines: the *classic* (but heavy) variability management chain



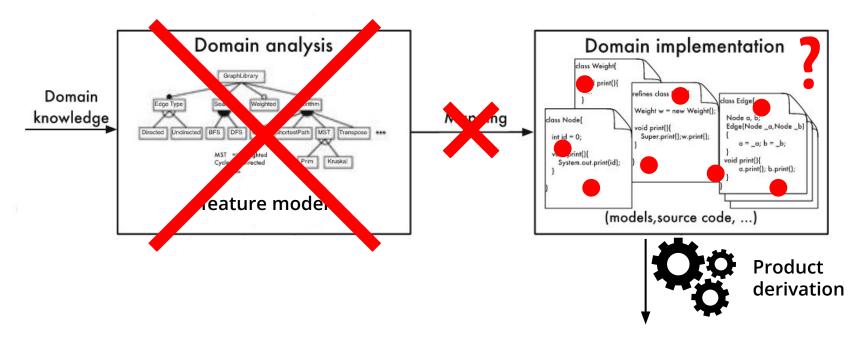
Source: http://stg-tud.github.io/sedc/Lecture/ws16-17/6-SPL.pdf

The reality of variability management



Source: http://stg-tud.github.io/sedc/Lecture/ws16-17/6-SPL.pdf

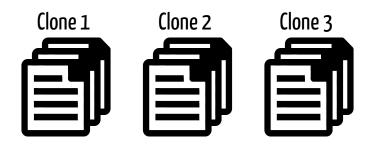
Problem: How to identify variability implementations in an existing OO codebase?



Source: http://stg-tud.github.io/sedc/Lecture/ws16-17/6-SPL.pdf

Feature location and feature identification: challenges and impact

<u>Context: projects clones</u>

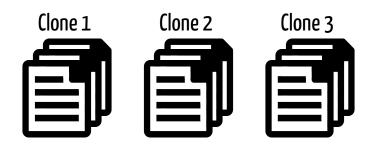


Detection method:

Comparison between clones and mapping with the domain features [Assunção2017]

Feature location and feature identification: challenges and impact

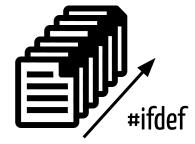
Context: projects clones



Context: unique codebase and

preprocessing directives

#ifdef \longrightarrow variant



Detection method:

Comparison between clones and mapping with the domain features [Assunção2017]

Detection method:

Determining the consistency of directives [Liebig2010]

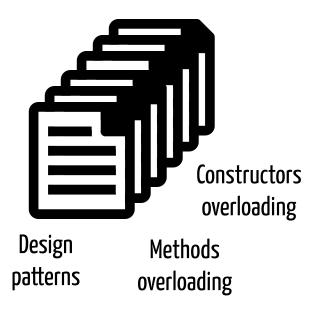
Feature location and feature identification: challenges and impact

<u>Our context: large and unique object-oriented codebase</u>

- Several implementation mechanisms
- Variability buried in the code (variation points)

Detection method:

- Techniques coupling static and dynamic analysis [Michelon2021]
- No technique using only static analysis
 [Metzger2014, Tërnava2017]



Inheritance

Features, variation points and variants

```
1 public abstract class Shape {
2 public abstract double area();
3 public abstract double perimeter(); /*...*/
4 }
```

```
public class Circle extends Shape {
5
      private final double radius;
6
      // Constructor omitted
7
      public double area() {
8
        return Math.PI * Math.pow(radius, 2);
9
10
      public double perimeter() {
11
        return 2 * Math.PI * radius;
12
13
14
```

```
public class Rectangle extends Shape {
15
      private final double width, length;
16
      // Constructor omitted
17
      public double area() {
18
         return width * length;
19
       3
20
       public double perimeter() {
21
         return 2 * (width + length);
22
23
       }
       public void draw(int x, int y) {
24
      // rectangle at (x, y, width, length)
25
26
       }
       public void draw(Point p) {
27
      // rectangle at (p.x, p.y, width, length)
28
       }
29
30
```

Features, variation points and variants

Inheritance

v_rectangle

```
vp_shape
    public abstract class Shape {
1
      public abstract double area();
2
      public abstract double perimeter(); /*...*/
3
4
                                               v_circle
    public class Circle extends Shape
5
      private final double radius;
6
      // Constructor omitted
7
      public double area() {
8
        return Math.PI * Math.pow(radius, 2);
9
       }
10
      public double perimeter() {
11
        return 2 * Math.PI * radius;
12
13
14
```

15	<pre>public class Rectangle extends Shape {</pre>
16	<pre>private final double width, length;</pre>
17	// Constructor omitted
18	<pre>public double area() {</pre>
19	<pre>return width * length;</pre>
20	}
21	<pre>public double perimeter() {</pre>
22	<pre>return 2 * (width + length);</pre>
23	}
24	<pre>public void draw(int x, int y) {</pre>
25	<pre>// rectangle at (x, y, width, length)</pre>
26	}
27	<pre>public void draw(Point p) {</pre>
28	<pre>// rectangle at (p.x, p.y, width, length)</pre>
29	}
30	}

Features, variation points and variants

1

Inheritance

v_rectangle

1	public abstract class Shape {
2	<pre>public abstract double area();</pre>
3	<pre>public abstract double perimeter(); /**/</pre>
4	1
4	5
	while along Cincle extends Shape (v_circle
5	public class circle extends snape {
6	<pre>private final double radius;</pre>
7	// Constructor omitted
8	<pre>public double area() {</pre>
9	<pre>return Math.PI * Math.pow(radius, 2);</pre>
10	}
11	<pre>public double perimeter() {</pre>
12	<pre>return 2 * Math.PI * radius;</pre>
13	}
14	}

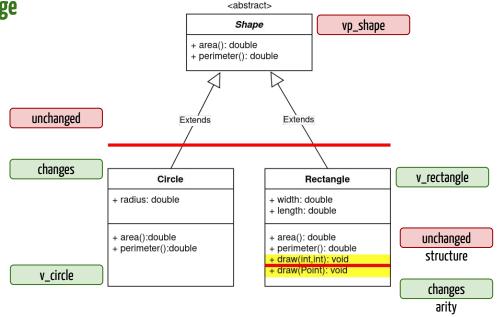
15	<pre>public class Rectangle extends Shape {</pre>
16	<pre>private final double width, length;</pre>
17	// Constructor omitted
18	<pre>public double area() {</pre>
19	<pre>return width * length;</pre>
20	}
21	<pre>public double perimeter() {</pre>
22	return 2 * (width + length);
23	} vp_draw
24	<pre>public void draw(int x, int y) {</pre>
25	<pre>// rectangle at (x, y, width, length)</pre>
26	}
27	<pre>public void draw(Point p) {</pre>
28	<pre>// rectangle at (p.x, p.y, width, length)</pre>
29	}
30	}
	Overloading

Symmetries in nature, human-made artifacts, and OO constructs

Symmetry represents **immunity** to a possible **change**

and is present in object-oriented constructs

including the ones implementing variability!





Some references on symmetry in OO constructs: [Coplien2000, Zhao2003]

Types of symmetries in OO constructs

	Class as type (Not identified by symfinder)	Class subtyping	Method / constructor overriding	Method / constructor overloading	Strategy	Factory	Decorator	Template
Commonality (unchange) ⇒vp	Туре	Superclass / Interface	Signature	Name	Superclass	Abstract creator and product	Components and decorator interfaces	Method defining the template
Variability (change) ⇒ variant	Objects	Subclasses / implementing classes	Implementa- tions in subclasses	Signature	Algorithms	Concrete creators and products	Concrete components and decorators	Steps used in the template

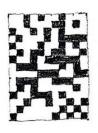
The theory of centres [Alexander2002]

Centre: a field of organized force in an object or part of an object which makes that object or part exhibit centrality.

A **centre** is commonly formed by one or multiple **local symmetry(ies)**.

 \Rightarrow The centre is the common part of the symmetric variants.

 $\underset{\longrightarrow}{\text{Random}}$ hard to describe



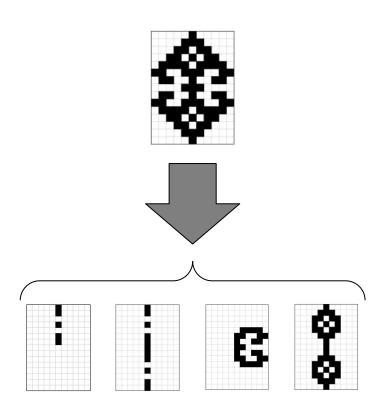


Ordered around a centre of symmetry → easy to describe

Centres and density

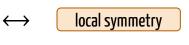
Local symmetries form a **structure**, whose **coherence** is determined with its **number of symmetries** [Alexander1968]

⇒ remarkable structures aggregate a
density of local symmetries



Identifying variability implementations

Variability implementation technique



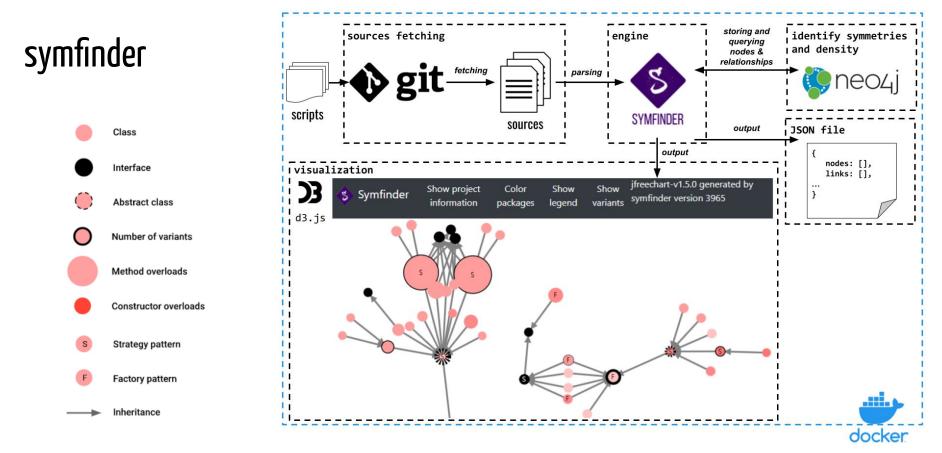
- variation point (commonality)
- variant (variability)



Identification through local symmetries in core assets

High density of symmetries \rightarrow variability intense places

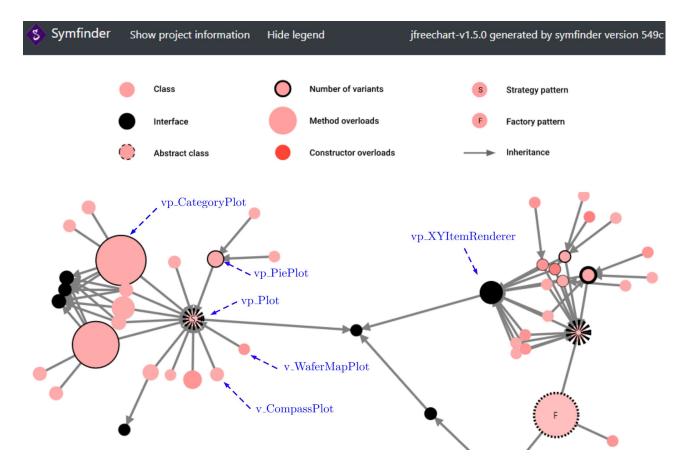
Xhevahire Tërnava, Johann Mortara, and Philippe Collet. 2019. Identifying and Visualizing Variability in Object-Oriented Variability-Rich Systems. In 23rd International Systems and Software Product Line Conference - Volume A (SPLC '19), September 9–13, 2019, Paris, France. ACM, New York, NY, USA, 12 pages.



Johann Mortara, Xhevahire Tërnava, and Philippe Collet. "symfinder: A Toolchain for the Identification and Visualization of Object-Oriented Variability Implementations". In: the 23rd International Systems and Software Product Line Conference. Vol. B. Paris, France: ACM Press, Sept. 2019, pp. 5–8.

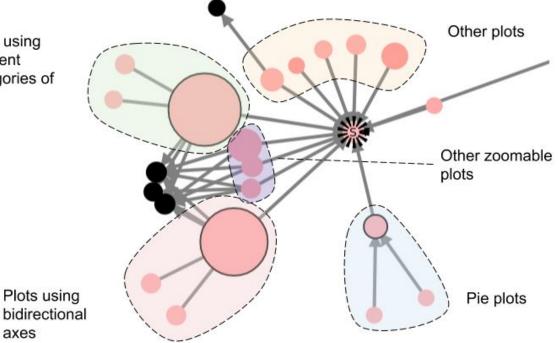
Johann Mortara, Philippe Collet, and Xhevahire Tërnava. "Identifying and Mapping Implemented Variabilities in Java and C++ Systems using symfinder". In: 24th ACM International Systems and Software Product Line Conference (SPLC '20). Ed. by ACM et al. Virtual Conference. MONTREAL, QC, Canada, Oct. 2020.

Visualization principles



Example of identified variability implementations

Plots using different categories of data

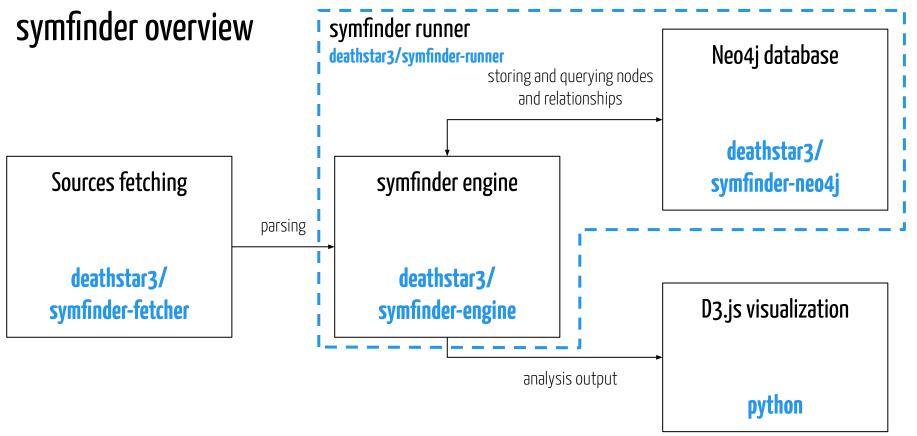


symfinder: first contact

Prerequisites

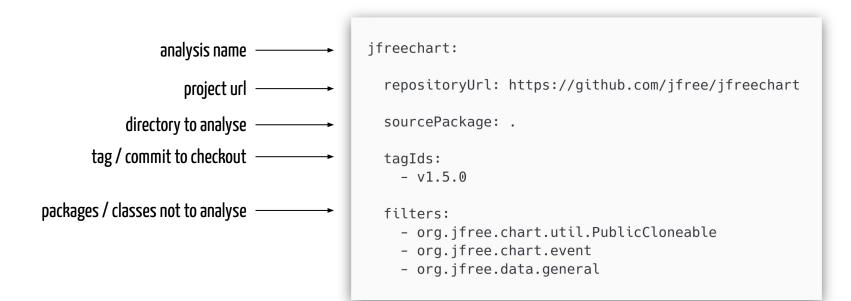
- Git to clone symfinder's repository, or ZIP download is also possible
- A functional Docker and Docker Compose install
 - Instructions → <u>https://docs.docker.com/engine/install/</u>
- A web browser to display the visualizations

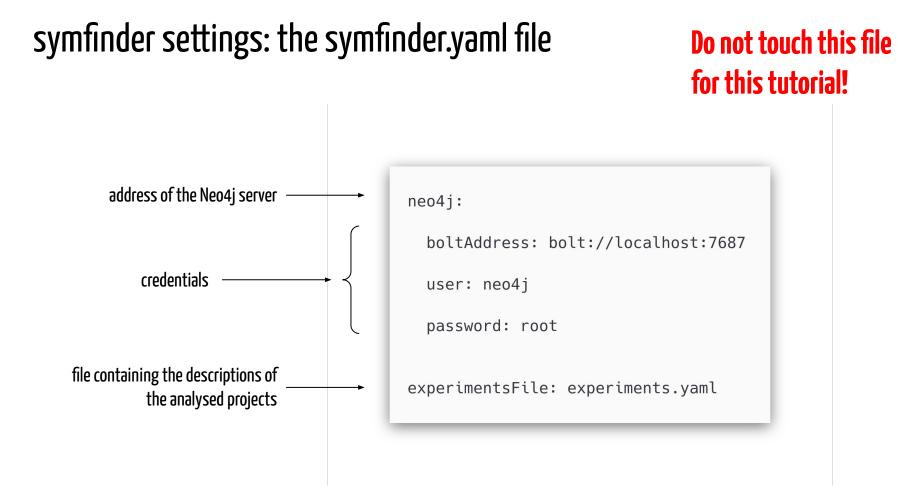
This list is also present in the REQUIREMENTS.md file at the repository's root.





Declaring studied projects: the experiments.yaml file





Running symfinder

Only one script to execute: run.sh

Parameters: projects to analyse

• • •

\$./run.sh jfreechart Creating resources directory Creating generated_visualizations directory Using splc2021-tuto images Cloning into 'resources/jfreechart'... Note: switching to 'tags/v1.5.0'.

You are in 'detached HEAD' state. You can look around, make experimental changes and commit them, and you can discard any commits you make in this state without impacting any branches by switching back to a branch.

If you want to create a new branch to retain commits you create, you may do so (now or later) by using -c with the switch command. Example:

git switch -c <new-branch-name>

Or undo this operation with:

git switch -

Turn off this advice by setting config variable advice.detachedHead to false

HEAD is now at fd72df7c Prepare for 1.5.0 release. HEAD is now at fd72df7c Prepare for 1.5.0 release.

Step 1: Fetching the sources

Docker images tag

Step 2: symfinder analysis

- the runner creates the containers
- symfinder waits for Neo4j to be started
- the analysis starts
 - visitors parse the code and store information in the Neo4j database
 - may take some time depending on the project's size!

.



End of analysis

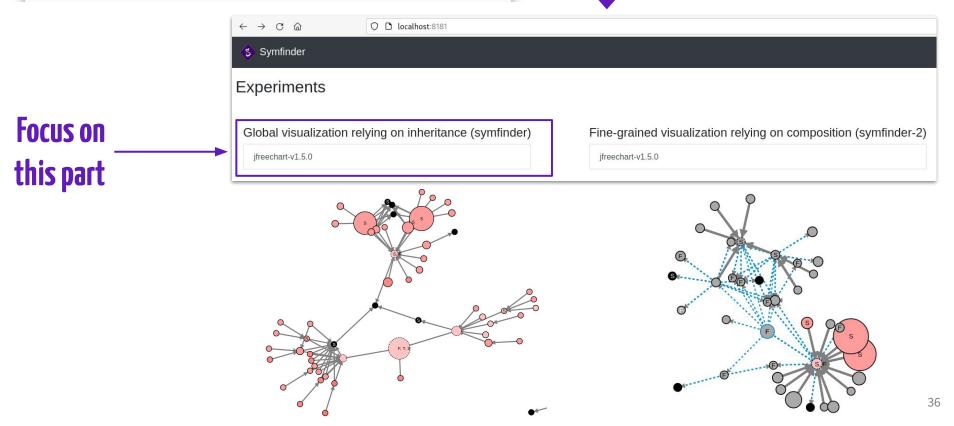
- after the visitors parsings are finished, the potential vp-s and variants are identified
- statistics on the analysis are displayed and symfinder stops
- the runner stops the containers

•••

symfinder-runner symfinder 09:21:19.953 [main] INFO visitors.SymfinderVisitor - Visitor: visitors.	
symfinder-runner symfinder 09:21:20.224 [main] INFO visitors.SymfinderVisitor - Visitor: visitors.	.ComposeTypeVisitor - Class:
org.jfree.data.category.IntervalCategoryDataset	
symfinder-runner symfinder 09:21:20.299 [main] INFO visitors.SymfinderVisitor - Visitor: visitors.	.ComposeTypeVisitor - Class:
org.jfree.data.category.DefaultCategoryDataset	
symfinder-runner symfinder 09:21:20.642 [main] INFO visitors.SymfinderVisitor - Visitor: visitors.	.ComposeTypeVisitor - Class:
org.jfree.data.category.SlidingCategoryDataset	
symfinder-runner symfinder 09:21:20.830 [main] INFO visitors.SymfinderVisitor - Visitor: visitors.	.ComposeTypeVisitor - Class:
org.jfree.data.category.CategoryToPieDataset	
symfinder-runner symfinder 09:21:20.984 [main] INFO visitors.SymfinderVisitor - Visitor: visitors.	
symfinder-runner symfinder 09:21:20.985 [main] INFO visitors.SymfinderVisitor - Visitor: visitors.	
symfinder-runner symfinder 09:21:21.011 [main] INFO visitors.SymfinderVisitor - Visitor: visitors.	.ComposeTypeVisitor - Class:
org.jfree.data.category.DefaultIntervalCategoryDataset	$\overline{)}$
symfinder-runner symfinder 09:21:21.417 [main] MY_LEVEL Symfinder - visitors.ComposeTypeVisitor exe	ecution time: 00:02:17.476
symfinder-runner symfinder 09:21:27.381 [main] MY_LEVEL Symfinder - Number of VPs: 926	
symfinder-runner symfinder 09:21:27.385 [main] MY_LEVEL Symfinder - Number of methods VPs: 454	
symfinder-runner symfinder 09:21:27.388 [main] MY_LEVEL Symfinder - Number of constructors VPs: 21	
symfinder-runner symfinder 09:21:27.395 [main] MY_LEVEL Symfinder - Number of method level VPs: 66	
symfinder-runner symfinder 09:21:27.396 [main] MY_LEVEL Symfinder - Number of class level VPs: 259	vp-s and variants
symfinder-runner symfinder 09:21:27.427 [main] MY_LEVEL Symfinder - Number of variants: 1923	
symfinder-runner symfinder 09:21:27.430 [main] MY_LEVEL Symfinder - Number of methods variants: 100	
symfinder-runner symfinder 09:21:27.433 [main] MY_LEVEL Symfinder - Number of constructors variants	
symfinder-runner symfinder 09:21:27.439 [main] MY_LEVEL Symfinder - Number of method level variants	
symfinder-runner symfinder 09:21:27.440 [main] MY_LEVEL Symfinder - Number of class level variants:	analysis are displayed and
symfinder-runner symfinder 09:21:27.452 [main] MY_LEVEL Symfinder - Number of nodes: 9526	
symfinder-runner symfinder 09:21:27.461 [main] MY_LEVEL Symfinder - Number of relationships: 11438	
symfinder-runner symfinder 09:21:27.470 [main] MY_LEVEL Symfinder - Number of corrected inheritance	e relationships: 265/14/8
symfinder-runner symfinder Jul 30, 2021 9:21:28 AM org.neo4j.driver.internal.logging.JULogger info	
symfinder-runner symfinder INFO: Closing driver instance 1489933928	
symfinder-runner symfinder Jul 30, 2021 9:21:28 AM org.neo4j.driver.internal.logging.JULogger info	
symfinder-runner symfinder INFO: Closing connection pool towards neo4j:7687	
symfinder-runner symfinder 09:21:28.701 [main] MY_LEVEL Symfinder - Total execution time: 00:05:23	.198
symfinder-runner symfinder exited with code 0	
symfinder-runner Stopping symfinder-neo4j	
symfinder-runner Stopping symfinder-neo4j done	
symfinder-runner Aborting on container exit	
symfinder-runner Removing symfinder symfinder-runner Removing symfinder-noodi	End of analysis
symmetrider - runner Removerig symmetrider - neo4j	
symfinder-runner Removing symfinder done	J
symfinder-runner Removing symfinder-neo4j done	
)	

\$./visualization.sh Creating network "symfinder-2_default" with the default driver Creating visualization ... done Attaching to visualization





First contact: running an existing experiment

Clone symfinder's repository and checkout the "tutorial" tag

Run symfinder on JFreeChart Display the visualization

~ \$ git clone https://github.com/DeathStar3/symfinder-SPLC2021-tutorial

~ \$ cd symfinder-SPLC2021-tutorial

~/symfinder-SPLC2021-tutorial \$ git checkout tutorial

~/symfinder-SPLC2021-tutorial \$./run.sh jfreechart

~/symfinder-SPLC2021-tutorial \$./visualization.sh

Repository's URL: <u>https://github.com/DeathStar3/symfinder-SPLC2021-tutorial/</u>

Requirements in the REQUIREMENTS.md file

Guided use of symfinder

Adapting symfinder for your project

1. Edit the **experiments/experiments.yaml** to add your project.

2. Relaunch symfinder on your project

~/symfinder \$./run.sh <project_name>



Hall of fame

The hall of fame of the tutorial will contain the visualizations you obtained on your projects!

If you wish to send us your visualizations:

- zip the **generated_visualizations** directory;
- send it to us by mail:
 - johann [dot] mortara [at] univ-cotedazur [dot] fr
 - philippe [dot] collet [at] univ-cotedazur [dot] fr

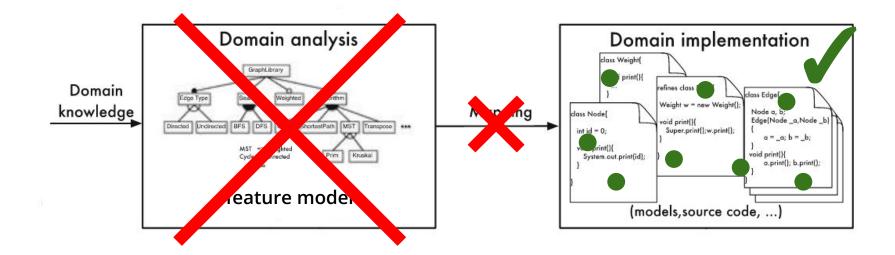
Wrapping up & Exchange time



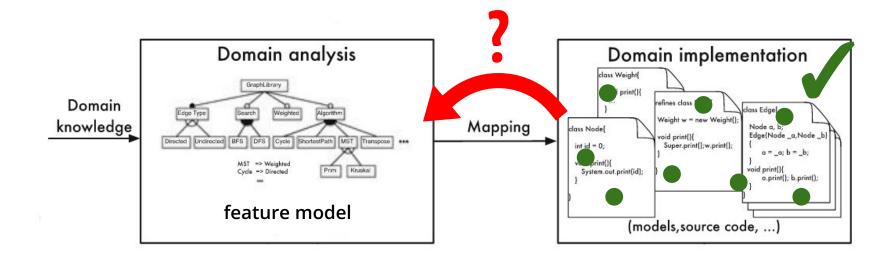
What you saw in this tutorial:

- Background on variability implementation in OO systems, and challenges around their identification
- Introduction on the notion of symmetry in 00 constructs
- The symfinder toolchain
 - how the toolchain uses density of symmetries to identify vp-s and variants in a single Java codebase
 - practical application on your systems

Potential variability implementations identified!

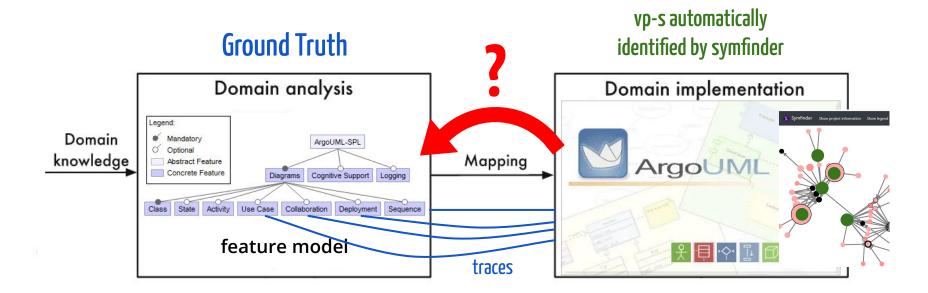


Potential variability implementations identified! but are they relevant w.r.t. existing features?



Source: http://stg-tud.github.io/sedc/Lecture/ws16-17/6-SPL.pdf

Application to a case study: ArgoUML



Question: Are the identified *vp-s* from ArgoUML relevant for a feature mapping?

Source: http://stg-tud.github.io/sedc/Lecture/ws16-17/6-SPL.pdf

Mapping vp-s and variants to features

Precision

Percentage of identified vp-s and variants that could be mapped to domain features

38%

- coarse grain features based on superficial domain knowledge
- not all identified places with a symmetry are related to variability

Recall

Percentage of features' traces that could be mapped to identified vp-s and variants

83%

The missing 17% of traces are not variability related (initialization classes, external libraries)

Johann Mortara, Xhevahire Tërnava, and Philippe Collet. "Mapping Features to Automatically Identified Object-Oriented Variability Implementations - The case of ArgoUML-SPL". In: 14th International Working Conference on Variability Modelling of Software-Intensive Systems (VaMoS '20). Magdeburg, Germany, Feb. 2020.

Mapping vp-s and variants to features

Precision

Percentage of identified vp-s and variants that could be mapped to domain features

38%

- coarse grain features based on superficial domain knowledge
- not all identified places with a symmetry are related to variability

Recall

Percentage of features' traces that could be mapped to identified vp-s and variants

83%

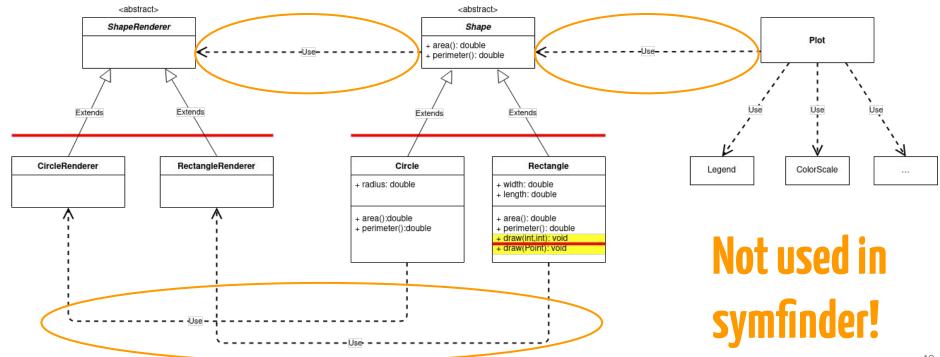
The missing 17% of traces are not variability

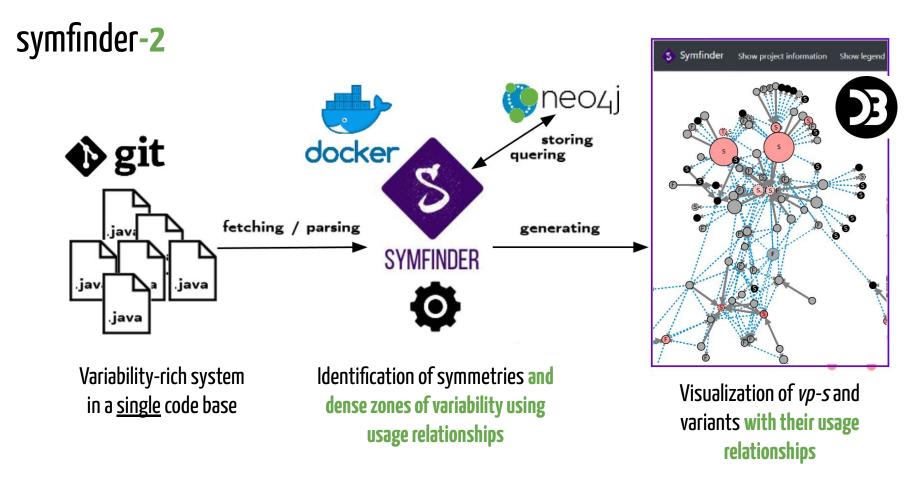
related (initialization classes, external libraries)

\Rightarrow need for a more precise detection method

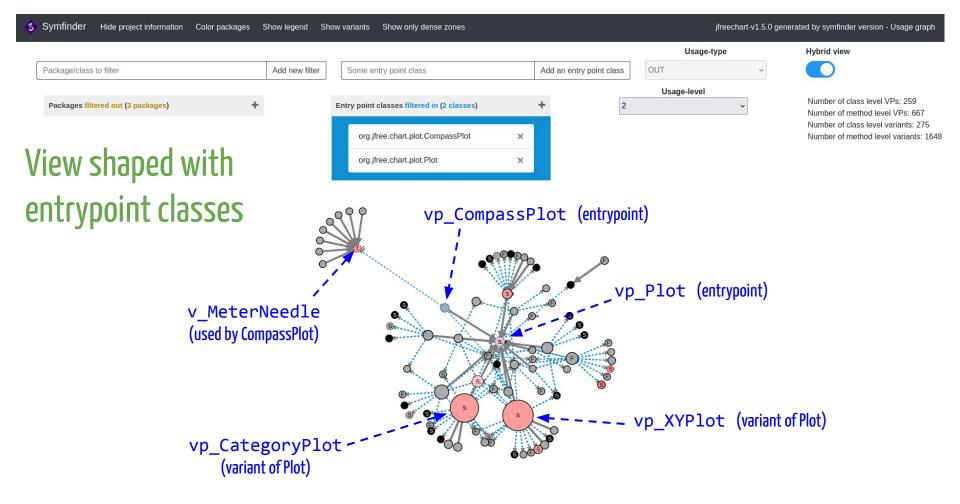
Johann Mortara, Xhevahire Tërnava, and Philippe Collet. "Mapping Features to Automatically Identified Object-Oriented Variability Implementations - The case of ArgoUML-SPL". In: 14th International Working Conference on Variability Modelling of Software-Intensive Systems (VaMoS '20). Magdeburg, Germany, Feb. 2020.

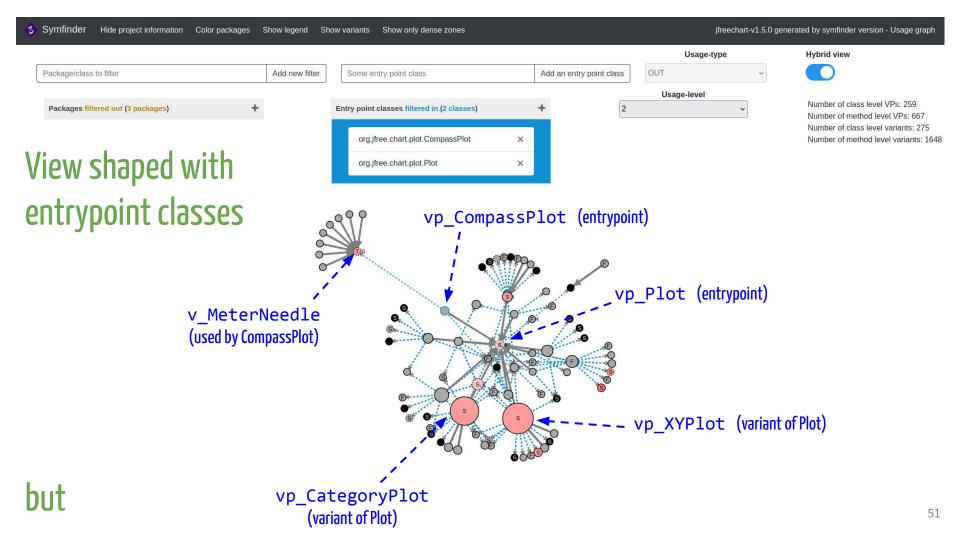
Extending the identification method



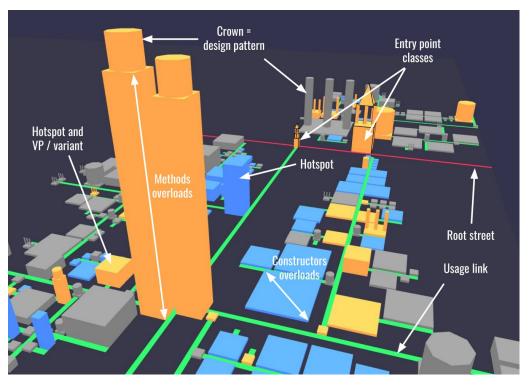


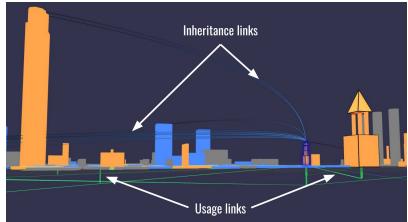
Johann Mortara, Xhevahire Tërnava, Philippe Collet, Anne-Marie Dery-Pinna. Extending the Identification of Object-Oriented Variability Implementations using Usage Relationships. SPLC 2021 - 25th ACM International Systems and Software Product Line Conference, Sep 2021, Leicester, United Kingdom. pp.1-8





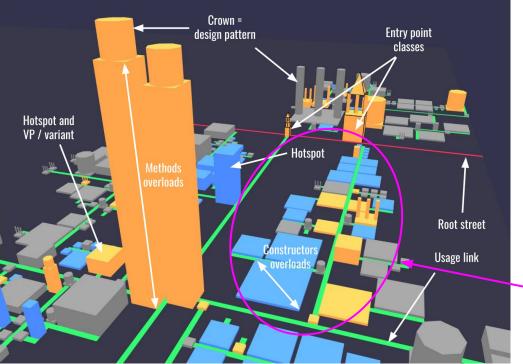
VariCity: visualizing variability implementations as a city

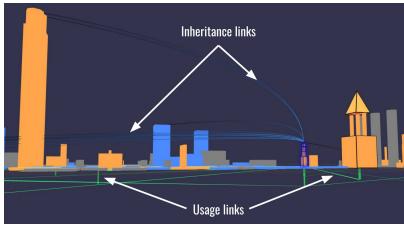




Johann Mortara, Philippe Collet, Anne-Marie Dery-Pinna. Visualization of Object-Oriented Variability Implementations as Cities. 9th IEEE Working Conference on Software Visualization (VISSOFT 2021), Sep 2021, Luxembourg (virtual), Luxembourg.

VariCity: visualizing variability implementations as a city





Density of variability implementations

Johann Mortara, Philippe Collet, Anne-Marie Dery-Pinna. Visualization of Object-Oriented Variability Implementations as Cities. 9th IEEE Working Conference on Software Visualization (VISSOFT 2021), Sep 2021, Luxembourg (virtual), Luxembourg.

Limitations / future work

Toolchain can analyze OO variability implementations projects in Java and C++

but no support for other OO languages (Python, JavaScript...) or implementation techniques

 \Rightarrow extend the toolchain support

Parameterized density gives first results but need to understand how to determine these parameters for a project

 \Rightarrow need to analyze projects with \neq architectures to understand better



Non exhaustive list of topics

- 1. Time of the conducted analyses on the different systems
- 2. Discussions on observed variability implementations and architectures
- 3. Discussion on the usage of the different features of the symfinder visualization
- 4. Feedback on the experience and possible improvements

Thanks to SPLC'21 sponsors!











How I Met Your Implemented Variability: Identification in Object-Oriented Systems with symfinder

Johann Mortara – Philippe Collet

Tutorial's website: <u>https://deathstar3.github.io/SPLC2021-symfinder-tutorial/</u>

Mail addresses to send your visualizations for the hall of fame:

- johann [dot] mortara [at] univ-cotedazur [dot] fr
- philippe [dot] collet [at] univ-cotedazur [dot] fr

Thank you for attending!

References

[Acher2018] Mathieu Acher. Software Variability and Artificial Intelligence. Ecole d'été du GDR GPL - EJCP 2018 <u>https://ejcp2018.sciencesconf.org/file/441457</u>

[Alexander1968] Christopher Alexander and Susan Carey. 1968. Subsymmetries. Perception & Psychophysics 4, 2 (1968), 73–77.

[Alexander2002] Christopher Alexander. 2002. The nature of order: an essay on the art of building and the nature of the universe. Book 1, The phenomenon of life. Center for Environmental Structure.

[Assunção2017] Wesley KG Assunção, Roberto E Lopez-Herrejon, Lukas Linsbauer, Silvia R Vergilio and Alexander Egyed. 2017. Reengineering legacy applications into software product lines: a systematic mapping. Empirical Software Engineering 22, 6 (2017), 2972–3016.

[Coplien2000] James O Coplien and Liping Zhao. 2000. Symmetry breaking in software patterns. In International Symposium on Generative and Component-Based Software Engineering. Springer, 37–54.

[Coplien2019] James O. Coplien and Liping Zhao. 2019. Toward a general formal foundation of design. Symmetry and broken symmetry. Technical Report. A VUB Lecture Series Publication. Working draft.

[Couto2011] Marcus Vinicius Couto, Marco Tulio Valente, and Eduardo Figueiredo. Extracting Software Product Lines: A Case Study Using Conditional Compilation. In 15th European Conference on Software Maintenance and Reengineering (CSMR), pages 191-200, 2011.

[Liebig2010] Jörg Liebig, Sven Apel, Christian Lengauer, Christian Kästner and Michael Schulze. 2010. An analysis of the variability in forty preprocessor-based software product lines. In Proceedings of the 32nd ACM/IEEE International Conference on Software Engineering-Volume 1. ACM, 105–114.

[Michelon2021] Gabriela K. Michelon, Lukas Linsbauer, Wesley K.G. Assunção, Stefan Fischer, and Alexander Egyed. 2021. A Hybrid Feature Location Technique for Re-engineering Single Systems into Software Product Lines. In 15th International Working Conference on Variability Modelling of Software-Intensive Systems (VaMoS'21). ACM, New York, NY, USA, Article 11, 1–9.

[Metzger2014] Andreas Metzger and Klaus Pohl. 2014. Software product line engineering and variability management: achievements and challenges. In Proceedings of the on Future of Software Engineering (FOSE 2014). ACM, New York, NY, USA, 70-84. DOI: <u>http://dx.doi.org/10.1145/2593882.2593888</u>

[Open2015] OpenSignal. Android Fragmentation Report. August 2015 <u>https://www.opensignal.com/sites/opensignal-com/files/data/reports/global/data-2015-08/2015_08_fragmentation_report.pdf</u>

[Tërnava2019] Xhevahire Tërnava, Johann Mortara, and Philippe Collet. 2019. Identifying and Visualizing Variability in Object-Oriented Variability-Rich Systems. In 23rd International Systems and Software Product Line Conference - Volume A (SPLC '19), September 9–13, 2019, Paris, France. ACM, New York, NY, USA, 12 pages.

[Tërnava2018] Xhevahire Tërnava and Philippe Collet. Identifying Variability Implementations with Local Symmetries. unpublished tech report. 2018.

[Tërnava2017] Xhevahire Tërnava and Philippe Collet. 2017. On the Diversity of Capturing Variability at the Implementation Level. In Proceedings of the 21st International Systems and Software Product Line Conference-Volume B. ACM, 81–88.

[Zhao2002] Liping Zhao and James O Coplien. 2002. Symmetry in class and type hierarchy. In Proceedings of the Fortieth International Conference on Tools Pacific: Objects for internet, mobile and embedded applications. Australian Computer Society, Inc., 181–189.

[Zhao2003] Liping Zhao and James Coplien. 2003. Understanding symmetry in object-oriented languages. Journal of Object Technology 2, 5 (2003), 123–134.