▲ Hochschule Harz
Hochschule für angewandte Wissenschaften

INF FAKULTÄT FÜR INFORMATIK

Finding Lost Features in Cloned Systems

Jacob Krüger, Louis Nell, Wolfram Fenske, Gunter Saake, Thomas Leich

Agenda

Motivation

Mapping Process

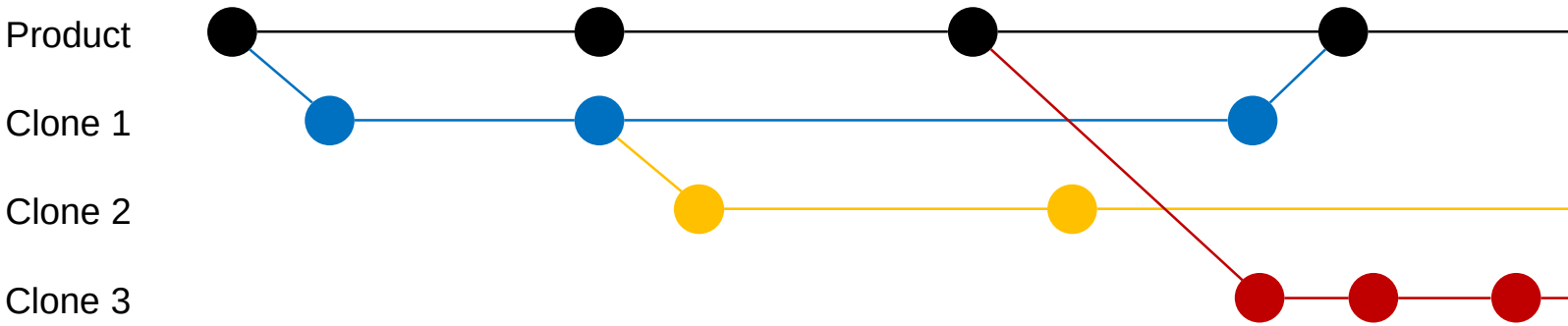
Evaluation

Conclusion

Software Reuse in Practice

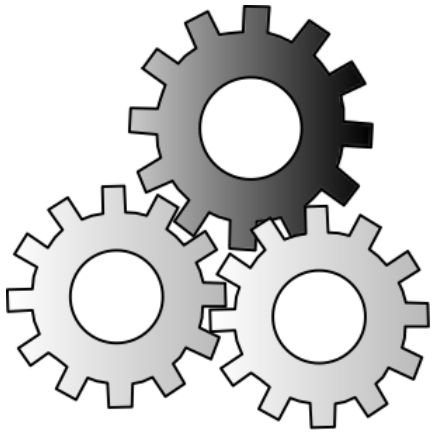
- Software reuse and customizing via cloning

- Potential problems?
  - Clones
  - Change propagation
  - Quality

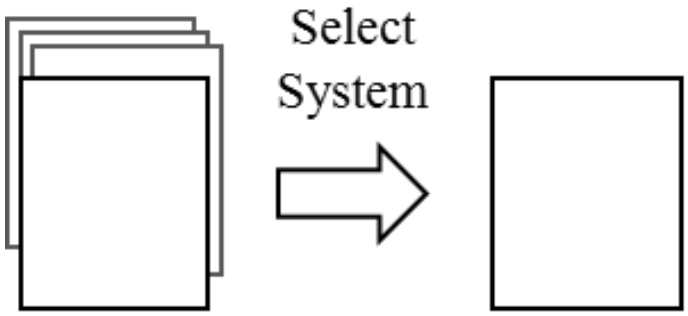


Improving the Situation

- Extract features into reusable artefacts / improve clone management
- What happens if knowledge about the clones is mising?
- Automated analysis:
  - Feature location
  - Code clone detection
  - …
- Still requires manual work!

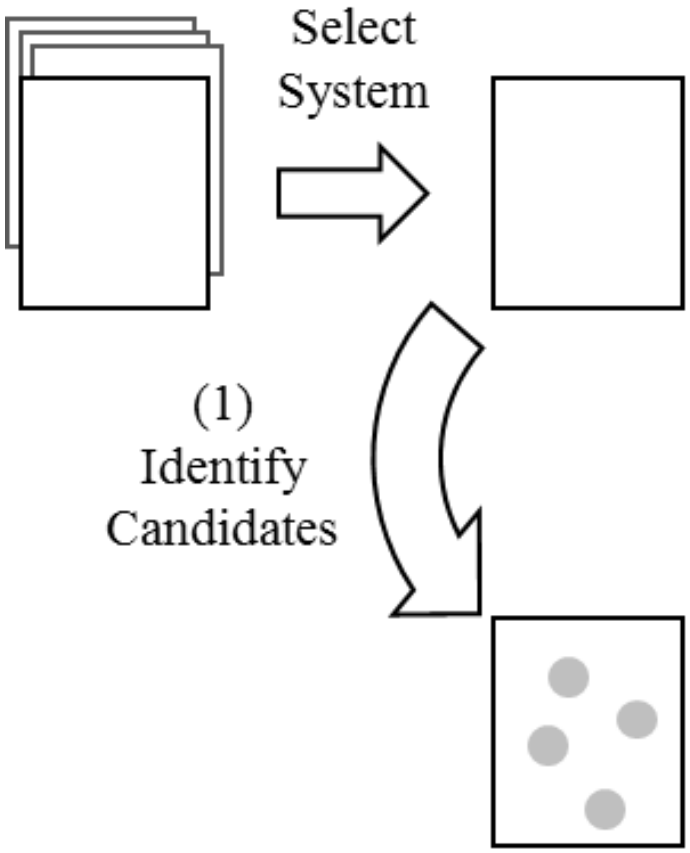


Mapping Process

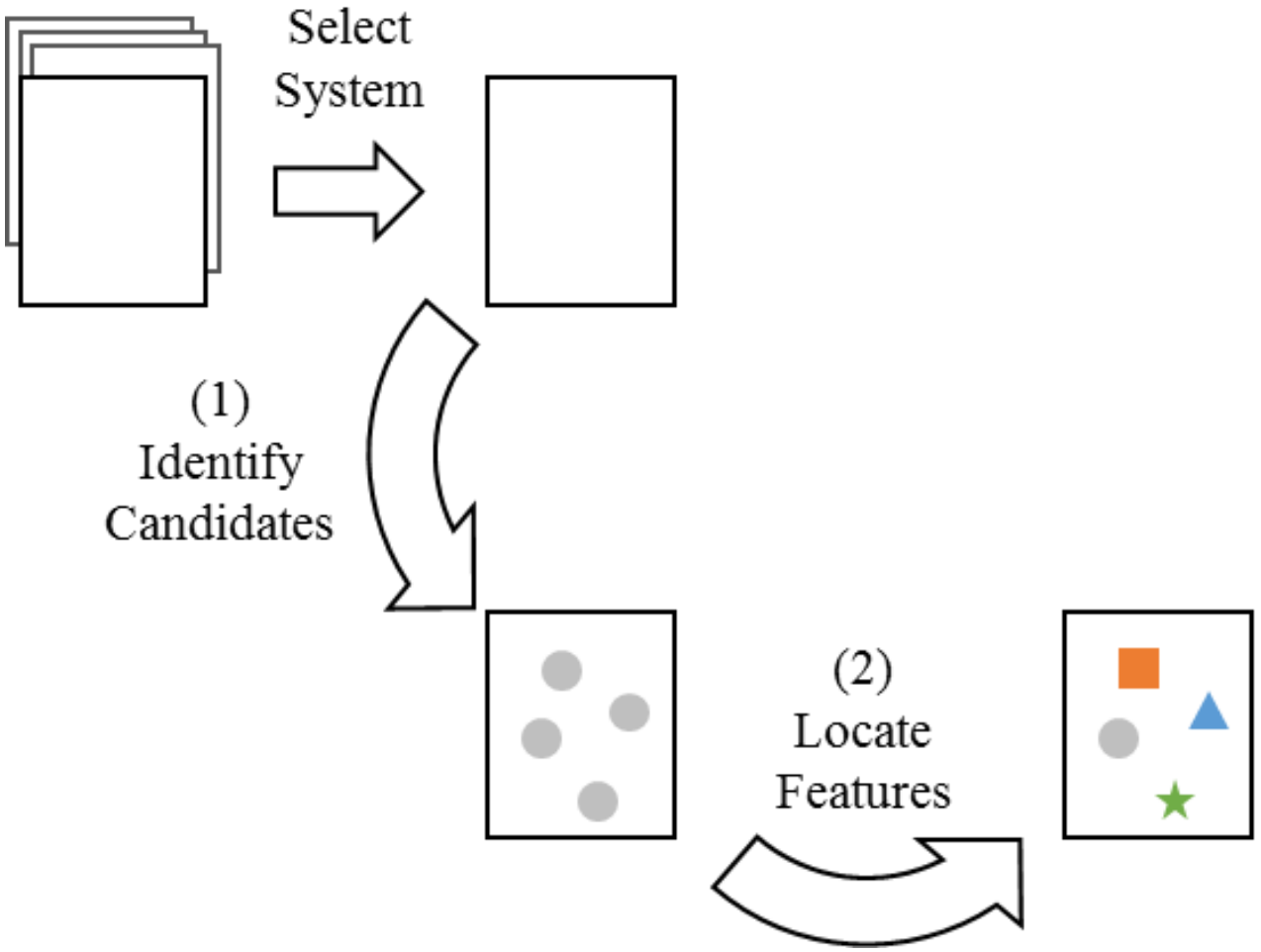


- Idea:
  - Language and tool independence
  - Step-wise to evaluate immediate results
  - Each considered clone analyzed once

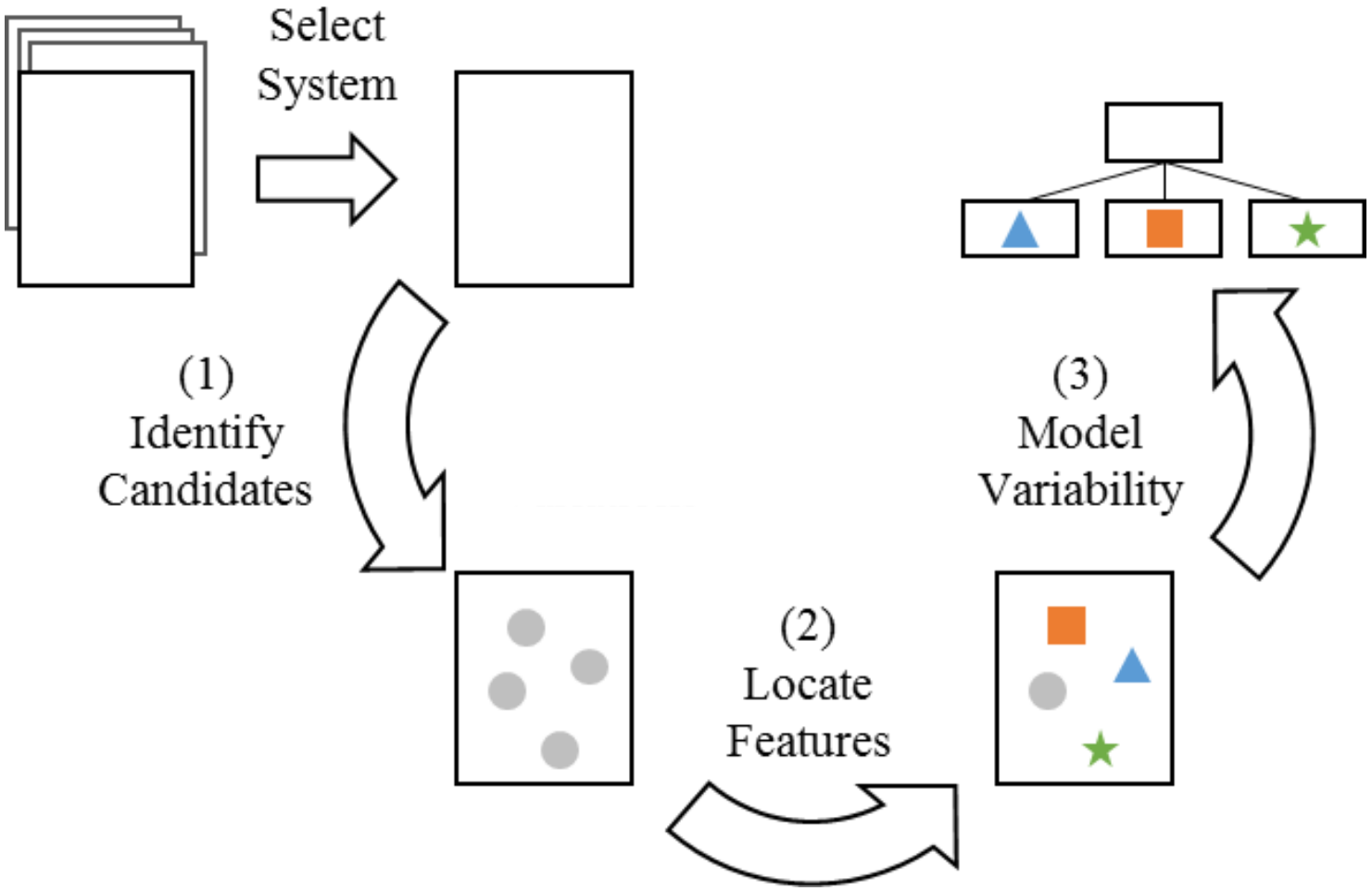
Mapping Process



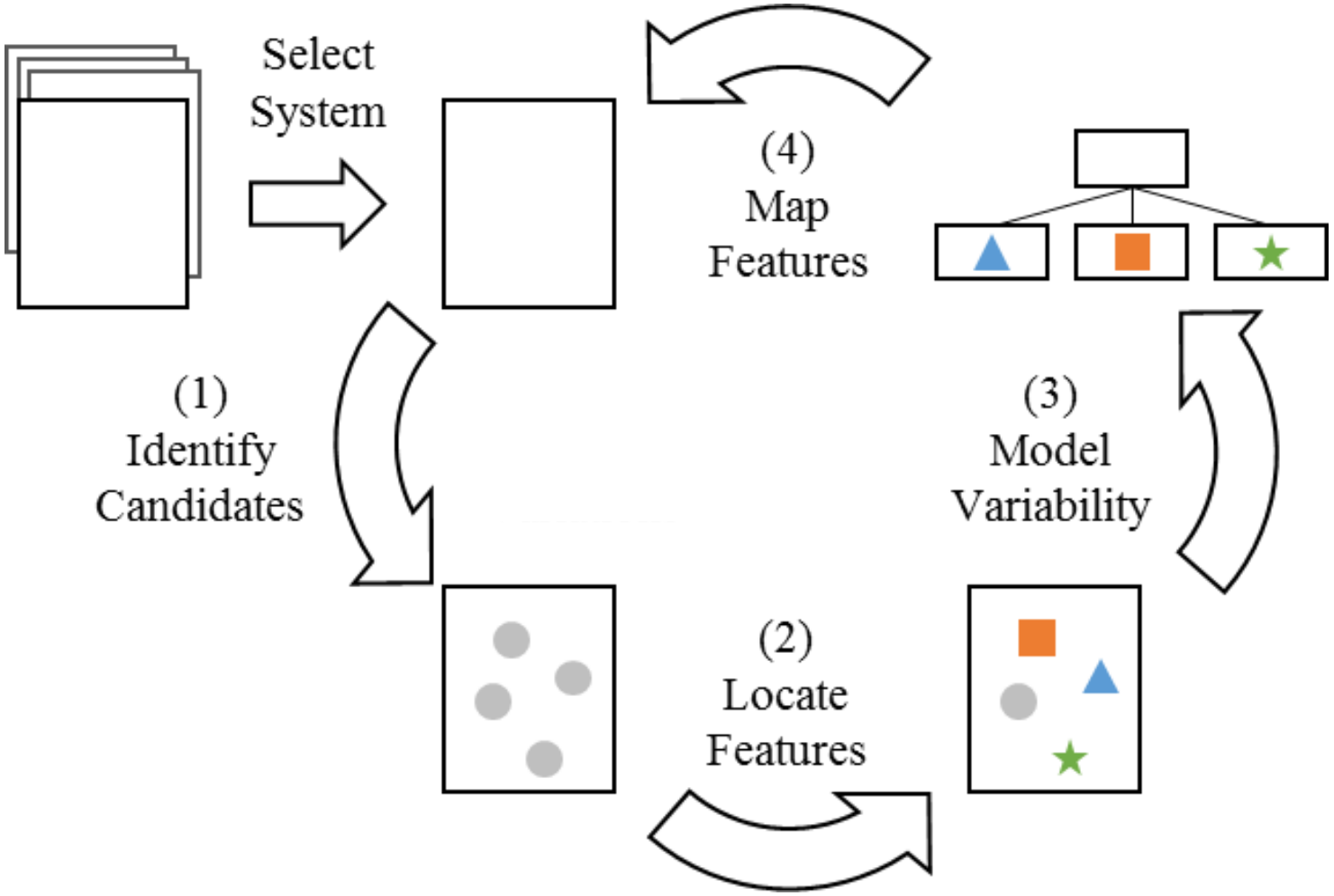
Mapping Process



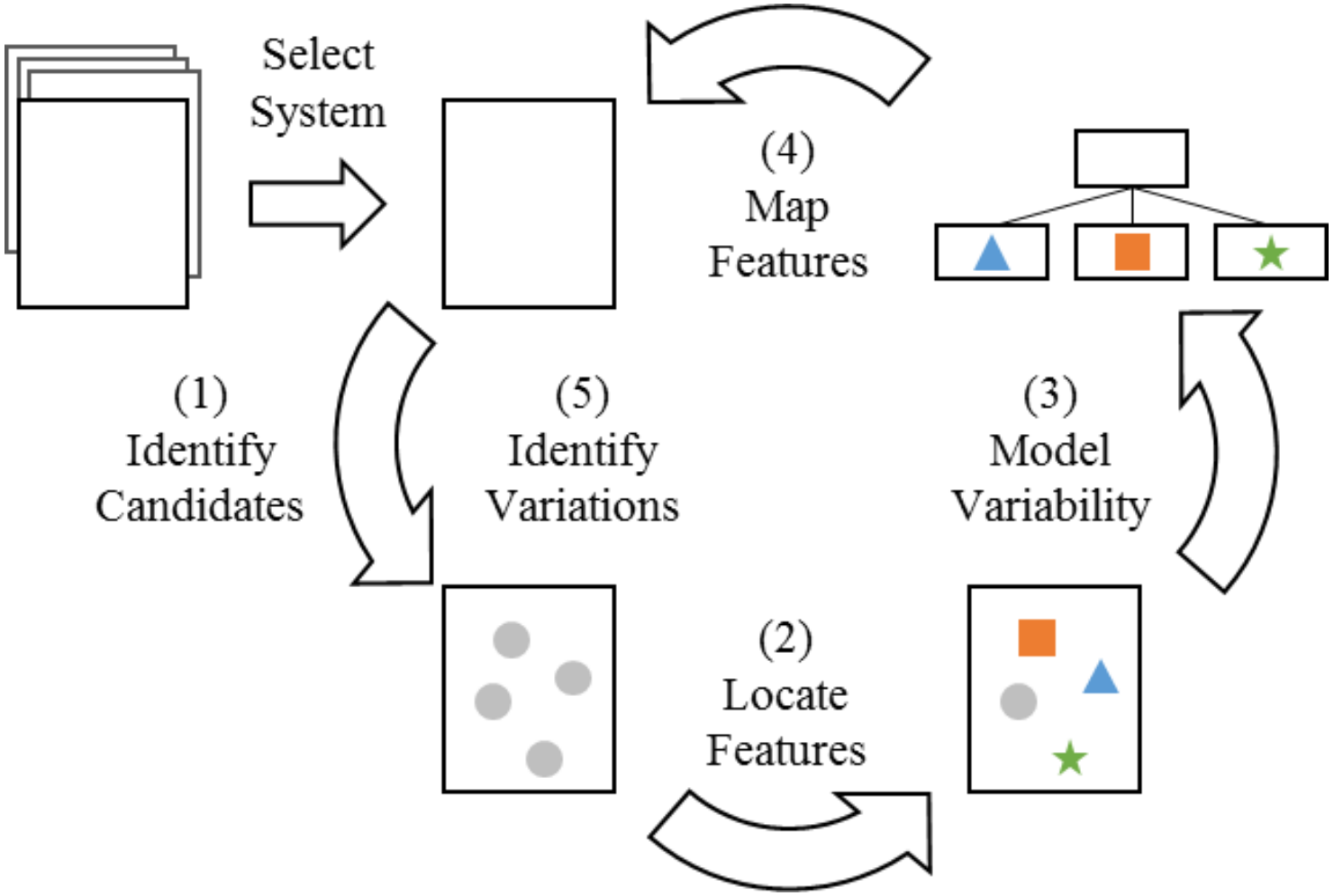
Mapping Process



Mapping Process

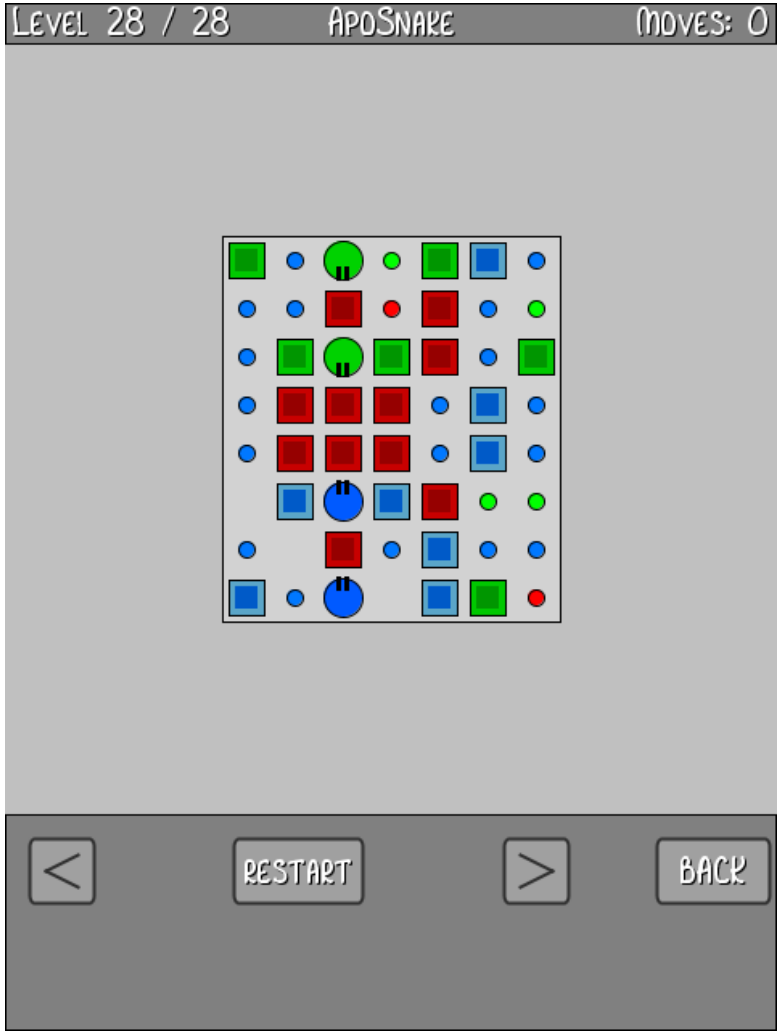


Mapping Process



Evaluation Setup

- Tools
  - Clone Detective
  - Visualization prototype
- Apo Games
  - Five cloned Java (Android) games
  - Reference study of one author [Fenske et al. SANER 2017]
- Used one clone as analysis system



Reference Study

- Completely automatic refactoring
  - Rename instance method
  - Pull-up to common feature

- 15 common and 5 unique features

System	Base #SLOC	Migrated #SLOC	ΔSLOC
Dice	2,504	1,346	1,158
Clock	3,584	2,696	888
MonoDroid	6,483	5,490	993
Snake	2,946	1,786	1,160
MyTreasure	5,322	4,483	839
Common	-	1,779	-
Total	20,839	17,580	-15.6%

Research Questions

RQ1 - How much of the extracted code in the reference study do we identify?

RQ2 - How do the differences between our mapping and the reference study look like?

RQ3 - How does a feature model for the subject systems look like?

Results

Feature	Function	#Classes	∑SLOC				
			Dice	Clock	MonoDroid	Snake	MyTreasure
Base	Core functionality	2	275	275	275	199	275
Editor	Create levels	1	61	34	0	27	0
Game	Game objects	8	347	129	118	253	20
Load	Load levels	2	282	282	282	282	267
Menu	Menu	2	295	61	201	93	0
Puzzle Chooser	Select levels	2	121	111	0	121	0
Save	Save levels	1	17	17	17	17	17
Total			1,398	909	893	992	579

RQ1 – Extracted Code

– In Dice and Clock we found more common code

– In Snake and MyTreasure we found less common code

➢ Manual analysis identified more common code in initial clone

➢ But expensive to identify common code among all clones

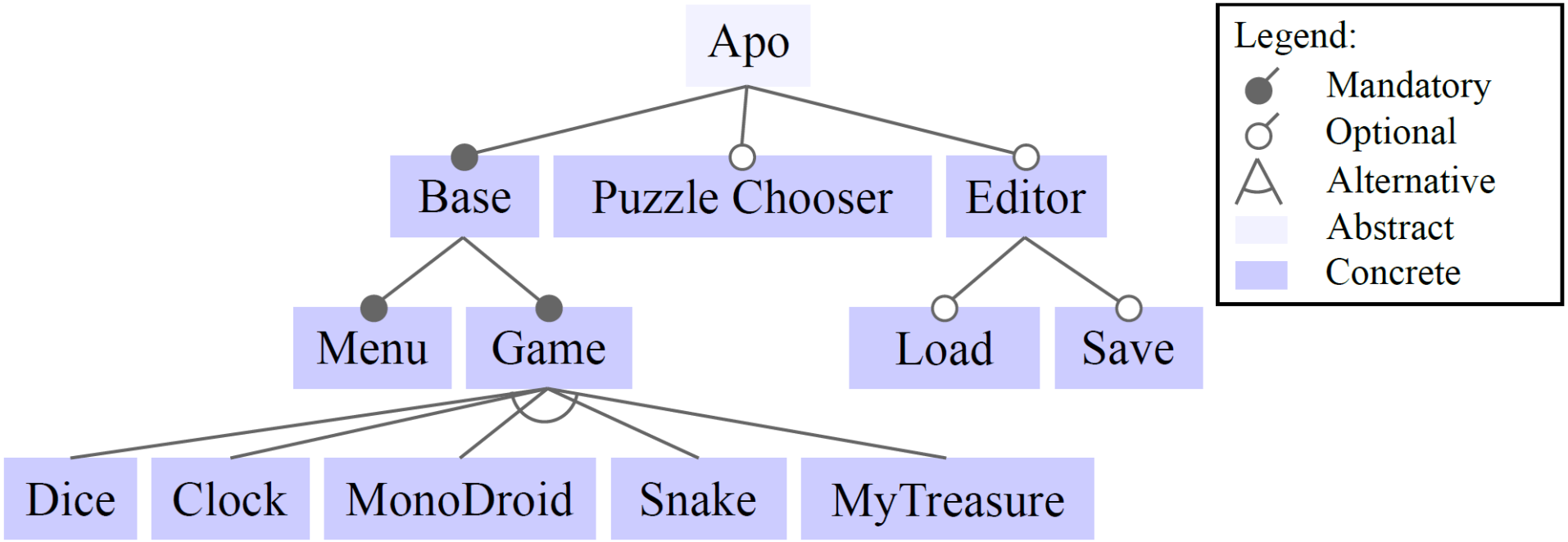
RQ2 – The Differences

- Manual analysis found more but smaller common code parts

- Automatic refactoring extracted common but also non-feature code

➢ Incomplete and scattered features with automatic refactoring
  - Automatic: 619 SLOC feature among all clones
  - Manual: ???

RQ3 – The Feature Model



Conclusion

– Analyzing features in cloned systems is an essential task …

– … but full automation seems unfeasible

– We proposed a semi-automatic process to guide developers

– Implementation and further evaluations in the future

▲ Hochschule Harz
Hochschule für angewandte Wissenschaften

OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG
INF FAKULTÄT FÜR INFORMATIK

Jacob Krüger
Phone +49 3943 – 659 878
E-Mail jkrueger @hs-harz.de|@ovgu.de
Friedrichstraße 57 – 59
38855 Wernigerode