Testing and Maintenance of Graphical User Interfaces

Valéria Lelli

PhD Thesis Defense, 19th November 2015 INSA/INRIA, Rennes – France

Jury

Lydie du BOUSQUET (Rapporteur) Philippe PALANQUE (Rapporteur) Pascale SÉBILLOT (Examinateur) Francois-Xavier DORMOY (Examinateur) Benoit BAUDRY (Directeur) Arnaud BLOUIN (Co-encadrant)



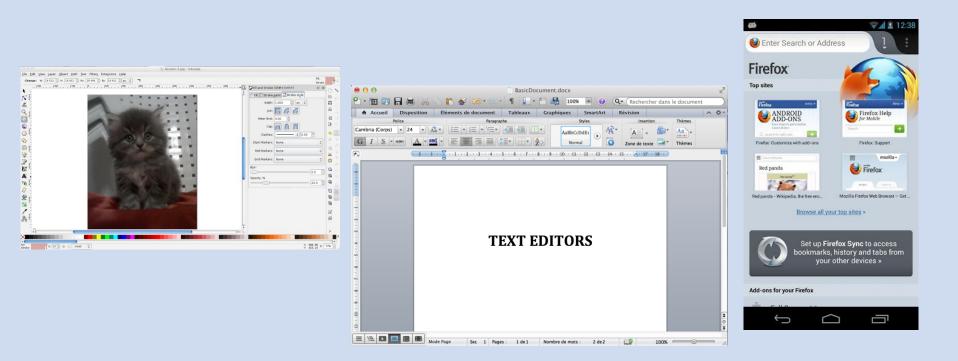




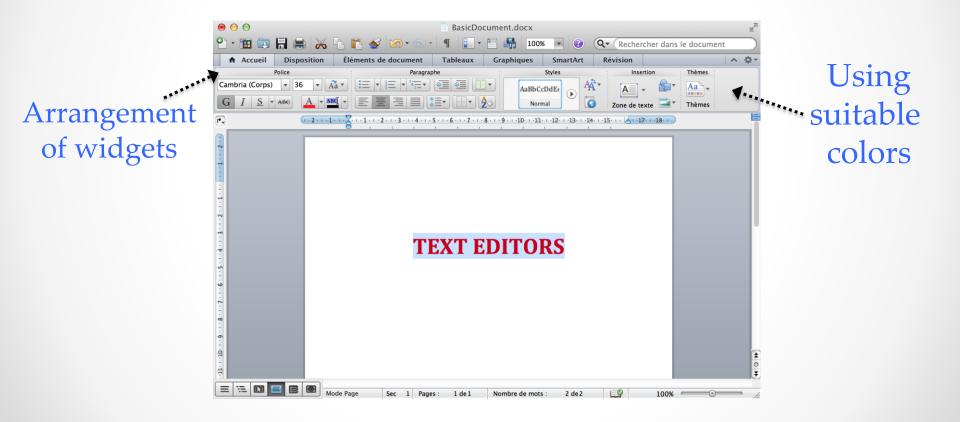


Graphical user interfaces (GUIs)

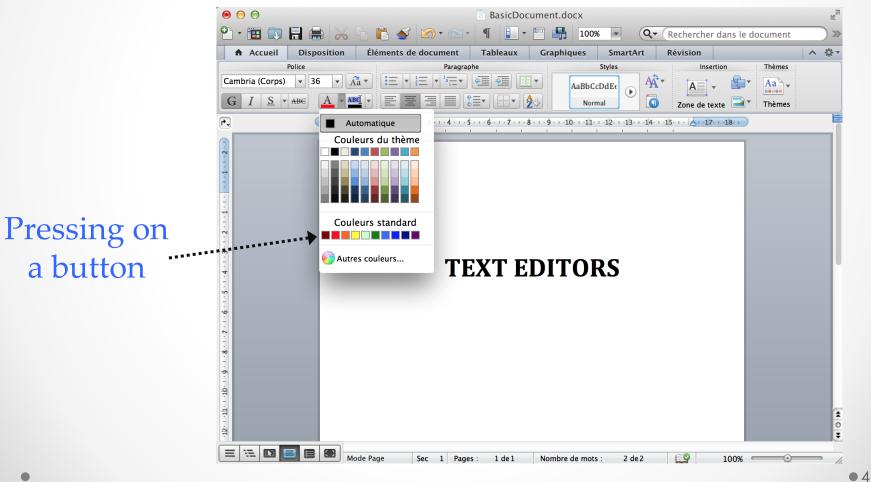
- Designed for being controlled by the users
- Composed of graphical interactive widgets



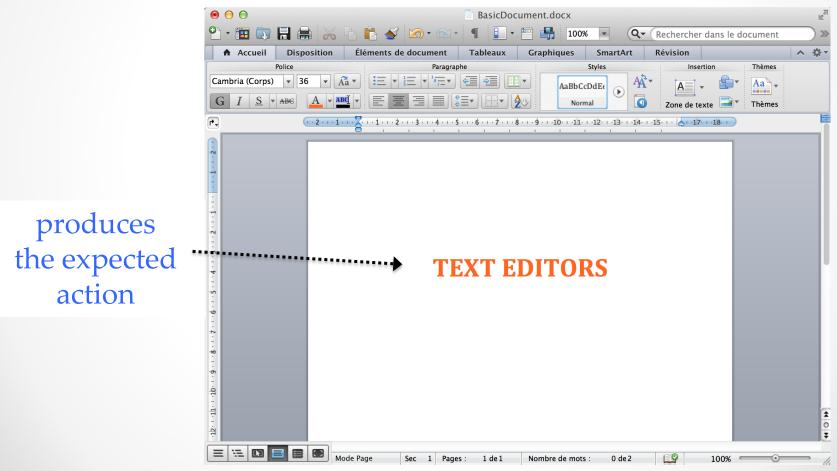
GUI designers concern the design and qualitative assessment of GUIs



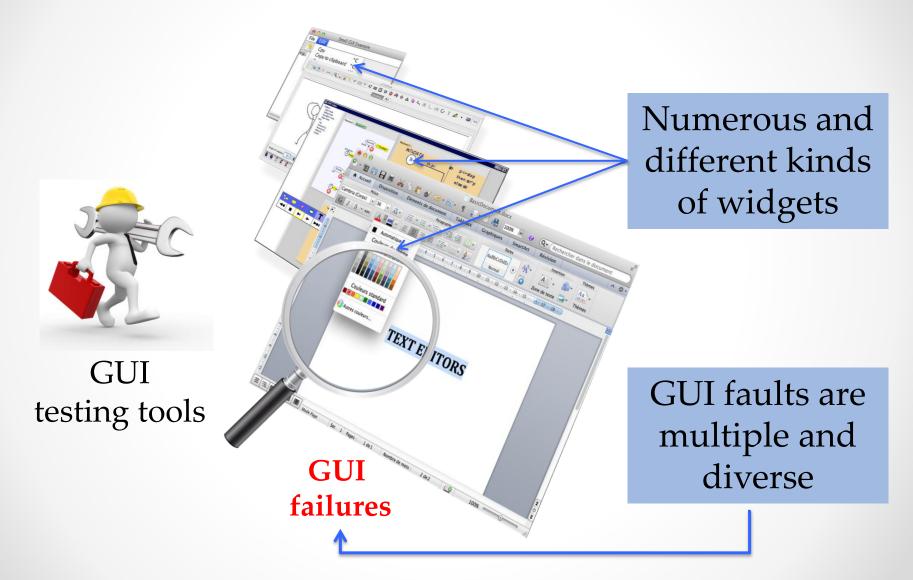
- Software engineers ensure that
 - GUIs react correctly to user interactions



- Software engineers ensure that
 - GUIs react correctly to user interactions



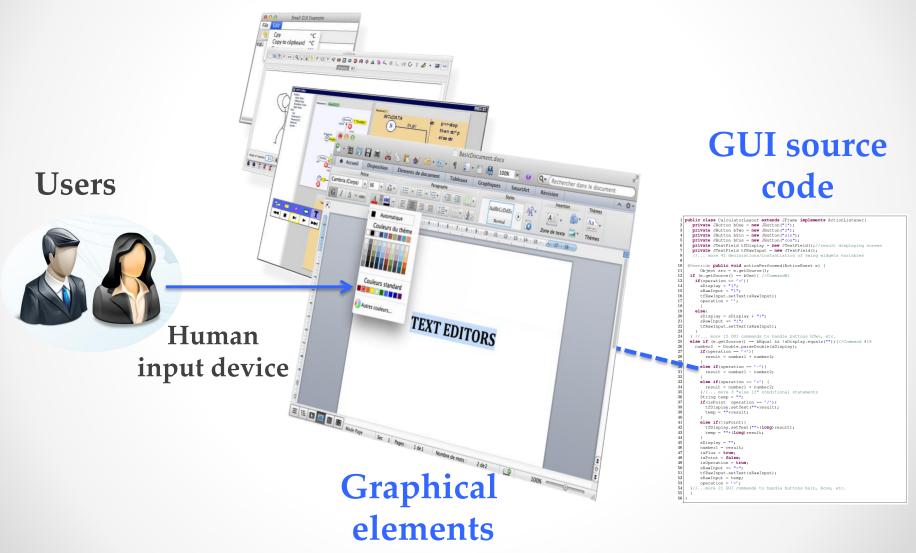
GUI testing

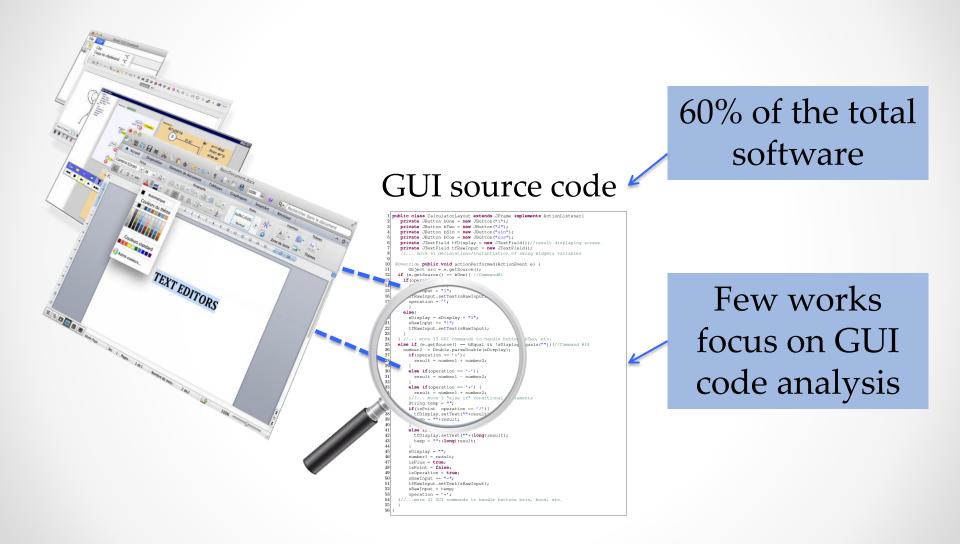


GUI testing techniques

- Capture and Replay tools
 - Recording user interactions to be replayed
- Monkey tools
 - Sending random events such as mouse events
- Functional GUI testing tools
 - Pre-defined libraries to write test cases
- Event-flow graphs
 - Based on the sequence of events to automatically generate test cases
 - ✗ GUI failures from the recent GUI developments

Graphical User Interfaces

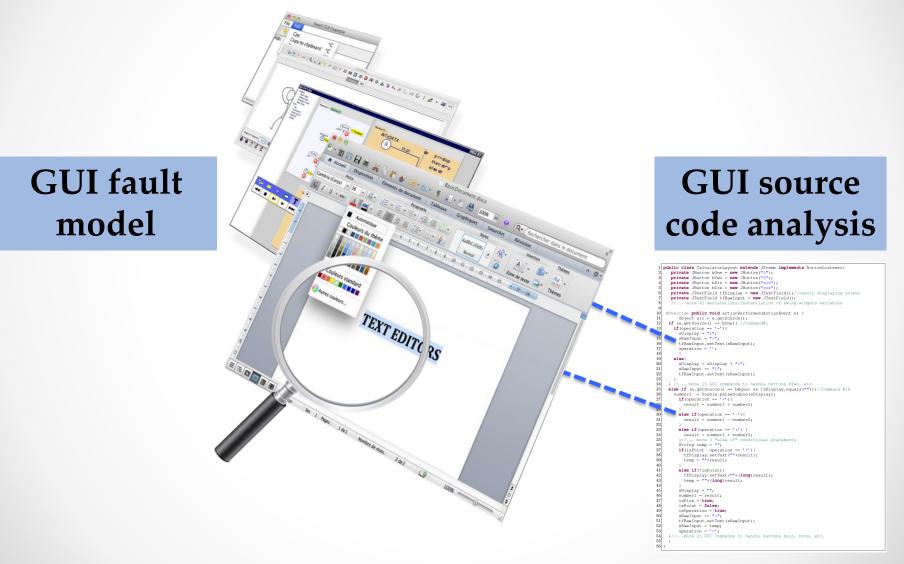




GUI code analysis

- GUI design smells
 - Bad coding practices that degrade GUI source code
- Bug finder tools
 - FindBugs and PMD do not focus on detect problems that affect the GUI source code
 - X Absence of GUI metrics/rules to detect GUI design smells

Contributions



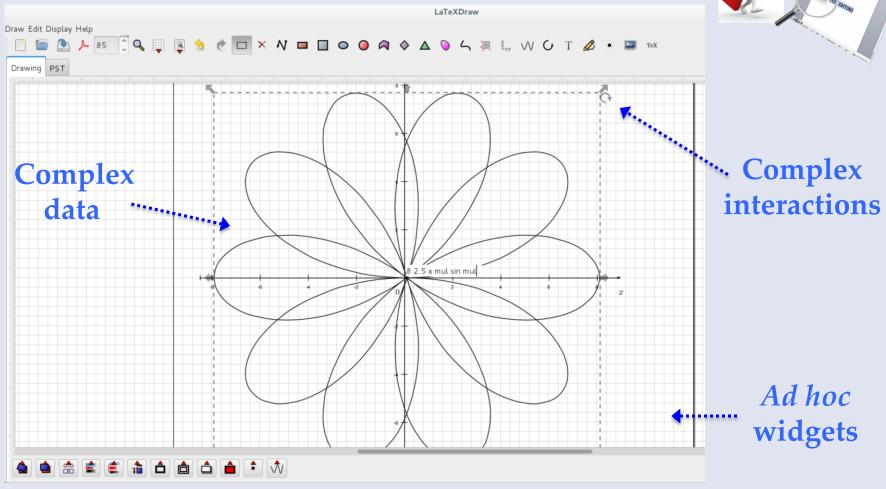


GUI Design



- Recent developments of GUIs involve more advanced user interactions
- Current GUI testing tools focus on finding bugs in classical GUIs
- How the characteristics of recent developments of GUIs impact on GUI testing?

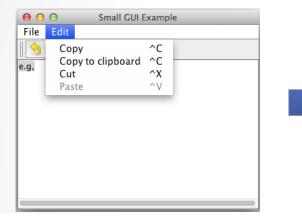
Post-WIMP GUIs



- *Ad hoc* widgets such as **drawing areas**
- Complex interactions: multi-touch, etc.

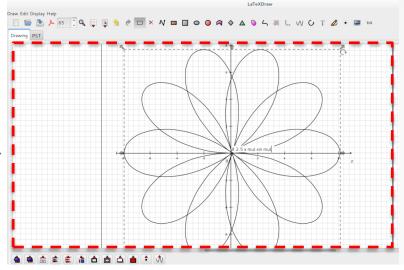
WIMP vs. post-WIMP GUIs





Event-based GUIs

- Standard widgets
- Mono-event interactions



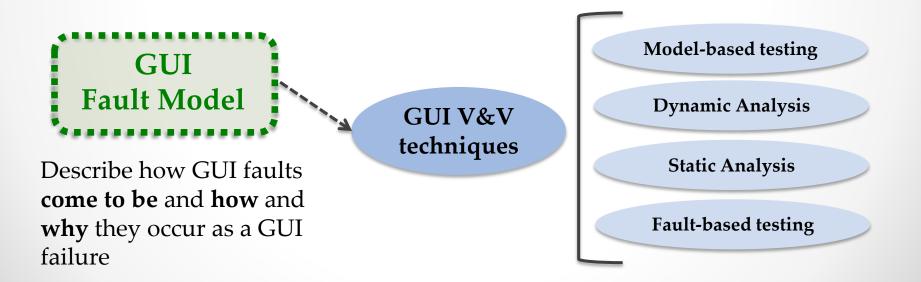
Interaction-based GUIs

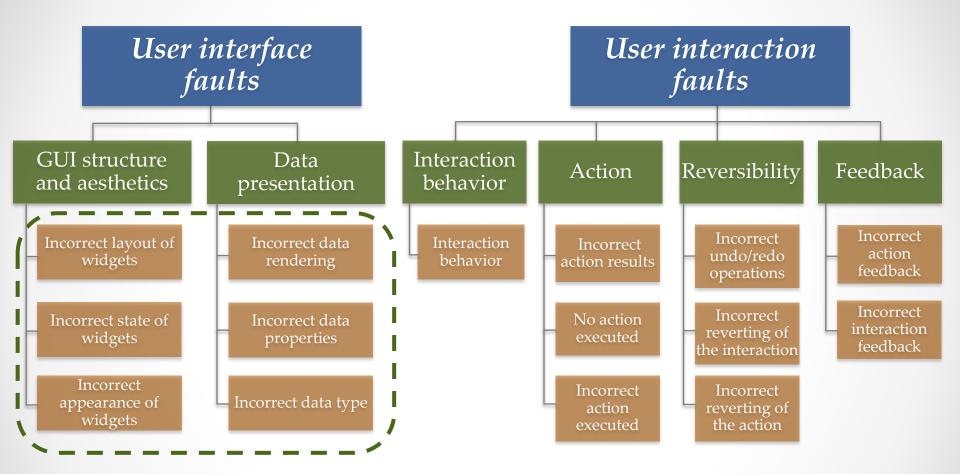
- Ad hoc widgets
- Multi-event interactions
- ✓ New problems of GUI faults



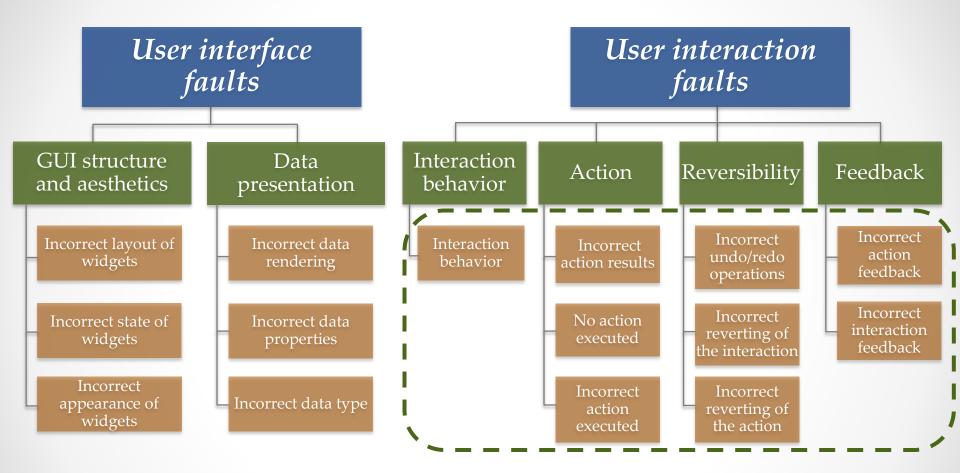
Current GUI testing tools

- Objectives
 - Baseline to evaluate the effectiveness of GUI testing techniques
 - Developing GUI testing techniques



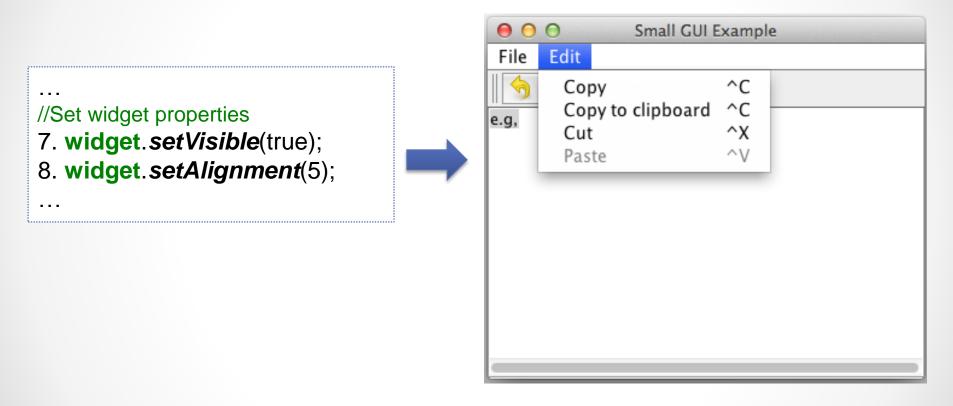


Structure and behavior of the graphical components



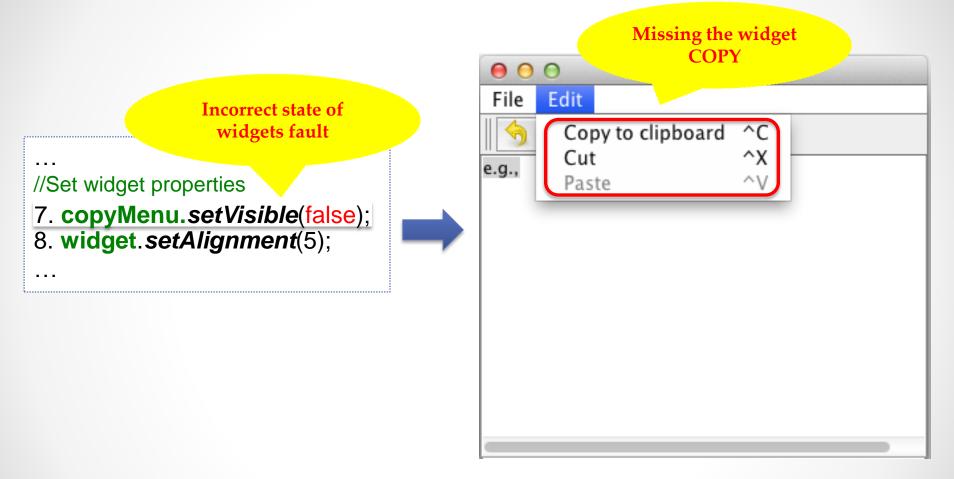
The interaction process when a user interacts with a GUI

Concrete examples of user interface faults



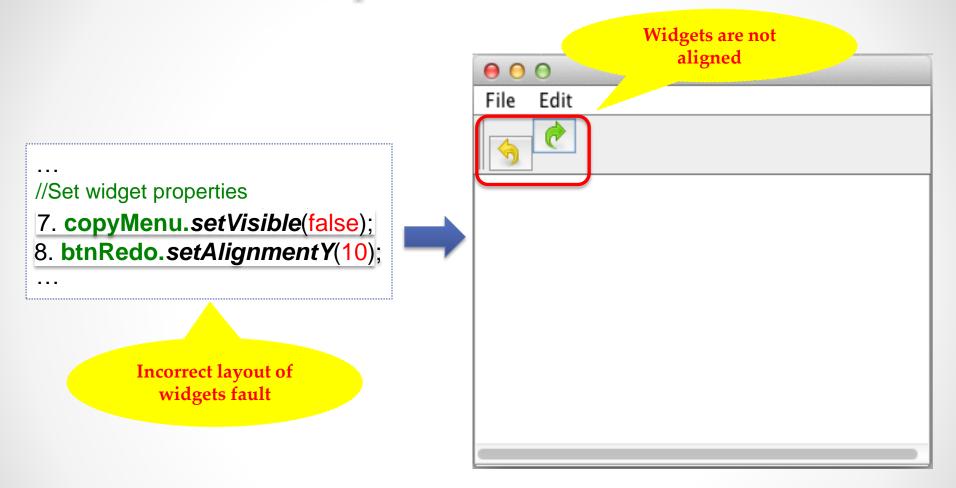
E.g. of GUI Fault: incorrect vs. correct lines of GUI code *E.g.* of GUI failure: a widget is not visible

Concrete examples of user interface faults



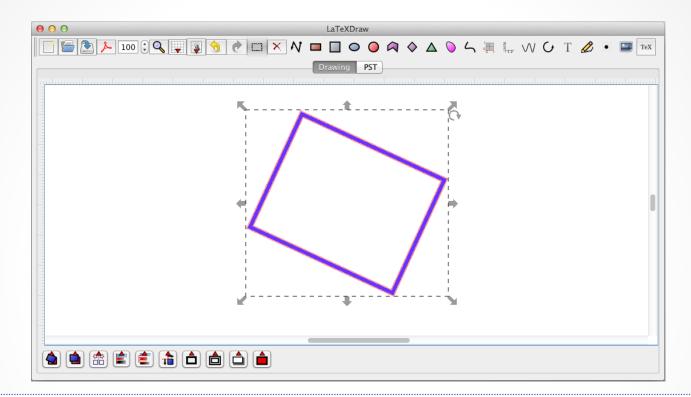
E.g. of GUI Fault: incorrect vs. correct lines of GUI code *E.g.* of GUI failure: a widget is not visible

Concrete examples of user interface faults



- *E.g.* of GUI Fault: incorrect vs. correct lines of GUI code *E.g.* of CUI failure: a widget is not visible
- *E.g.* of **GUI failure**: a widget is not visible

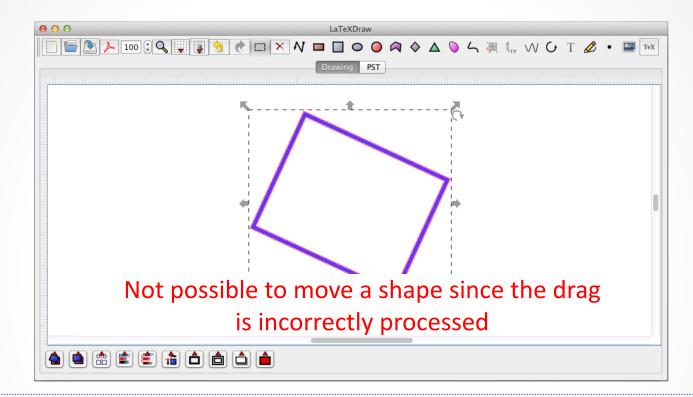
Concrete examples of user interaction faults



Interaction behavior

1. figures.firstElement().onDragged(formerPt, newPt);

Concrete examples of user interaction faults



Interaction behavior

1. figures.firstElement().onDragged(newPt, formerPt);

Fault model assessment

RQ1: Is the **GUI fault model** relevant against **real GUI failures**?

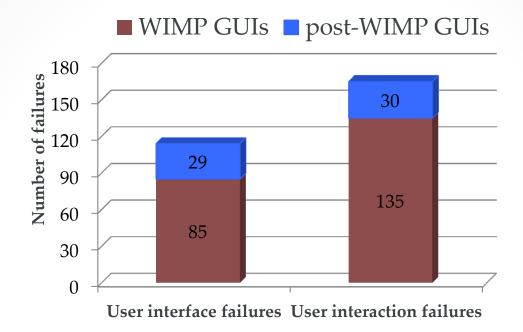
RQ2: Are **GUI testing tools** able to detect the **failures classified** in our fault model?

Experiment (RQ1): relevance

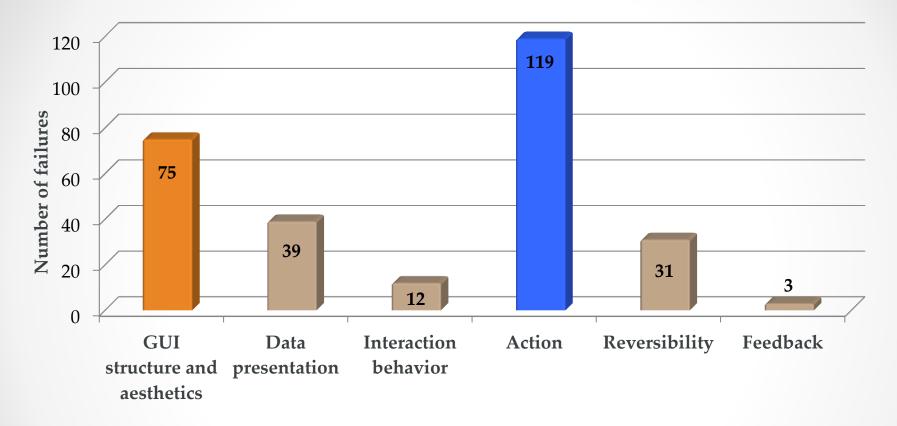
- GUI bug reports of 5 open-source software systems
 - Sweet Home 3D
 - File-roller
 - JabRef
 - Inkscape
 - Firefox Android

Several kinds of widgets, interactions and platforms

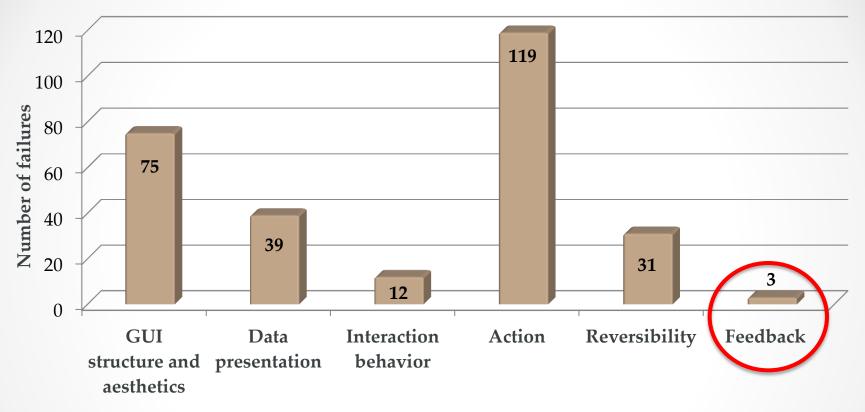
- Manual analysis of the real GUI bug reports
 - Source forge, bugzilla, etc.
 - Root cause: description, patches, comments, or stack traces



- All GUI failures (279) were classified into the fault model
- User interface (41%) and user interaction (59%)
- Post-WIMP
 - ✓ 25% (user interface)
 - ✓ 18% (user interaction)



- 43% into Action and
- 27% into GUI structure and aesthetics



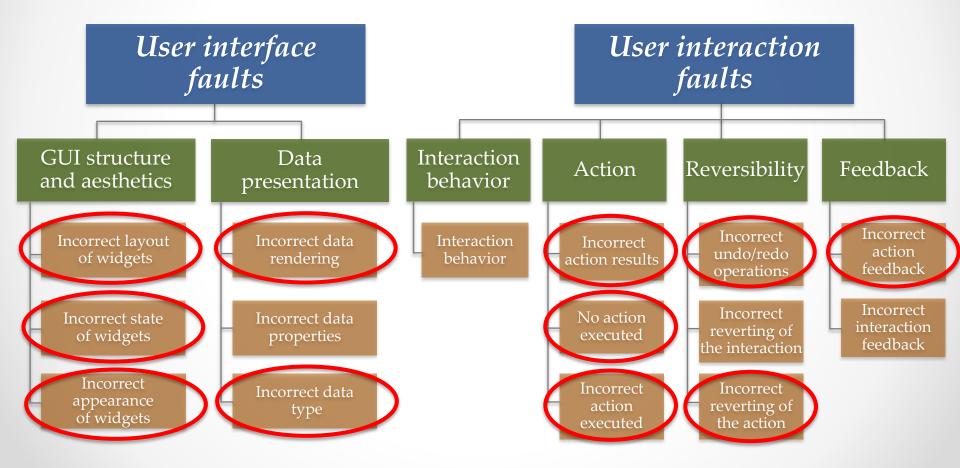
- 1% of GUI failures classified into Feedback
- Several "failures" were considered by developers as improvements

RQ1: Is the GUI fault model relevant against real GUI failures?

- ✓ All GUI related faults of 5 large scale case studies can be classified
 ✓ 279 GUI bug reports
- ✓ All the 6 fault categories are covered
- ✓ Faults concern WIMP and post-WIMP GUIs
 ✓ *Ad hoc* widgets (59 faults)

Experiment (RQ2): GUI testing tools

JabRef: selected 11 out of 15 GUI faults



Experiment (RQ2): GUI testing tools

- GUITAR²
 - Most popular academic tool in GUI testing
 - Automated test cases generation
 - Event-flow graph is built by reverse engineering
- Jubula³
 - Partially manual generation of test cases
 - Reuse pre-defined libraries to create manually test cases

 ²B. N. Nguyen, B. Robbins, I. Banerjee, A. Memon: GUITAR: an innovative tool for automated testing of GUI-driven software. Autom. Softw. Eng. 21(1): 65-105 (2014)
 ³http://www.eclipse.org/jubula

- GUITAR detected
 - 3 out of 11 GUI faults
- Jubula detected
 - 9 out of 11 GUI faults
- GUI failures detected
 ✓ Properties of standard widgets
 ✓ Crashes
 → Oracle for standard widgets

GUITAR

- Missed 8 out of 11 GUI faults reported in JabRef
- GUITAR builds the event-flow graph by
 - Extracting the sequence of events behind standard widgets
 - Collecting the information in the properties of standard widgets as event logs
 - X User interface failures into properties of standard widgets
 - X Complex data in *ad hoc widgets*
 - X Events are both widgets and their underlying interactions
 - X Ad hoc widgets and their multi-event interactions

Real example of a GUI failure

📕 JabRef File Edi	it View	BibTeX	Tools Options Help		✿ ①	
			🔗 Search	₩F		
18 🕞 🔒 🛷 🗈 🐧	Q	+	🔗 Incremental search 🛛 🗘	₩F		
		* 0		ЖR		
			Find duplicates	oon		
			Find and remove exact duplicates			
🔗 Search 🛛 🗶	#	Author	This and remove exact duplicates			
	1	Valéria	Manage content selectors		d Qualifying GUI defects	
	2	Valéria		ЖG	d Qualifying GUI Defects	
Search Clear	3	Yue, Ta			ansition from use cases to {UML} state machin	
	4	Yuan, X	Insert selected citations into LyX	ЖL	vent Sequence-Based Test Cases Using {GUI}	
💽 Incremental	5	Yuan, X	🥩 Insert selected citations into WinEdt 🛛 🏠	ЖW	ion Testing: Incorporating Event Context	
🔵 Highlight	6	Xie, Qin	💈 Fetch Medline	F5	d comparing automated test oracles for {GUI}-	
	/	van Dar			Jser interfaces	
🔵 Float	8	Takala,	🗇 Fetch CiteSeer	F6	of System-Level Model-Based (GUI) Testing of	
Settings 🦻	9 10	Silva, Jo Shneide	und bereit	F7	7 User Interface Testing With Spec Explorer an ulation: a step beyond programming language	
	10	Nguyen			testing of multiple GUI variants using the {GU	
	12	Nguyen	🔀 Open PDF or PS	F4	Inctionality testing through {GUIs}	
	13	Navarre	Open URL or DOI F3 M New subdatabase based on AUX file		lel-based user interface description techniqu	
	14	Myers, I			, and future of user interface software tools	
	15	Mori, Gi			evelopment of Multidevice User Interfaces thr	
	16	Memon.	Juntegrity check		Fault-Detection Effectiveness of {GUI} Test Cas	
	17		Aui M. and Sona, Mary Lou and Pollack, Mar Covera	iue cri	· · · · ·	
	18				esting of {GUIs}	
	19		eonardo and Pezz\`{e}, Mauro and Riganel AutoBla			
	20	Mariani, Leonardo and Pezz\`{e}, Mauro and Riganel AutoBlackTest: a tool for automatic black-box testing				
	21		, , Clemens Nylandsted and Beaudouin-Laf {VIGO}:			
	22	Kervinen	Antti and Maunumaa, Mika and Katara, Mika Model-	-Base	d Testing Through a {GUI}	
	23	Hackner,	Daniel R. and Memon, Atif M. Test ca	ase ge	nerator for {GUITAR}	
	24	Grechani	K, Mark and Xie, Qing and Fu, Chen Maintai	ining a	and evolving {GUI}-directed test scripts	
	25	Graham,	Dorothy and Fewster, Mark Experie	ences	of Test Automation: Case Studies of Software	
	26				rns: elements of reusable object-oriented soft	
	27	El-Far, Ib	rahim K. and Whittaker, James A. Model-	-base	d software testing	

Status: Marked selected entry

Real example of a GUI failure

 $\bigcirc \bigcirc \bigcirc$

🕐 🗁 📳

of 🗈 🔒 🤌

\$

References EICS paper*							
🔗 Search 🛛 🗙	e	ŧ	Author	Title			
	1		Valéria Lelli and Blouin, Arnaud and Baudry, Benoit	Classifying and Qualifying GUI defects			
	2		Valéria Lelli and Blouin, Arnaud and Baudry, Benoit	Classifying and Qualifying GUI Defects			
Search Clear Incremental Highlight Float Settings	3		Yue, Tao and Ali, Shaukat and Briand, Lionel	Automated transition from use cases to {UML} state machin			
	4		Yuan, Xun and Memon, Atif M.	Generating Event Sequence-Based Test Cases Using (GUI)			
	5		Yuan, Xun and Cohen, Myra B. and Memon, Atif M.	{GUI} Interaction Testing: Incorporating Event Context			
	6	;	Xie, Qing and Memon, Atif M.	Designing and comparing automated test oracles for {GUI}-			
	7	,	van Dam, Andries	Post-{WIMP} user interfaces			
	8		Takala, Tommi and Katara, Mika and Harty, Julian	Experiences of System-Level Model-Based {GUI} Testing of			
	9	1	Silva, Jos{\'e} L. and Campos, Jos{\'e} Creissac and P	Model-based User Interface Testing With Spec Explorer an			
	1	0	Shneiderman, Ben	Direct manipulation: a step beyond programming language			
	1	1	Nguyen, Duc Hoai and Strooper, Paul and Suess, Jorn	Model-based testing of multiple GUI variants using the {GU			
	1	2	Nguyen, Duc Hoai and Strooper, Paul and S\"{u}\ss, J	Automated functionality testing through {GUIs}			
	1	3	Navarre, David and Palanque, Philippe and Ladry, Je	{ICOs}: A model-based user interface description techniqu			
	1	.4	Myers, Brad and Hudson, Scott E. and Pausch, Randy	Past, present, and future of user interface software tools			
	1	5	Mori, Giulio and Paterno, Fabio and Santoro, Carmen	Design and Development of Multidevice User Interfaces thr			
	1	6	Memon, Atif M. and Xie, Qing	Studying the Fault-Detection Effectiveness of {GUI} Test Cas			
	1	7	Memon, Atif M. and Soffa, Mary Lou and Pollack, Mar	Coverage criteria for {GUI} testing			
	1	.8	Memon, Atif M. and Soffa, Mary Lou	Regression testing of {GUIs}			
	1	9	Mariani, Leonardo and Pezz\`{e}, Mauro and Riganel	AutoBlackTest: Automatic Black-Box Testing of Interactive			
	2	0	Mariani, Leonardo and Pezz\`{e}, Mauro and Riganel	AutoBlackTest: a tool for automatic black-box testing			
	2	1	Klokmose, , Clemens Nylandsted and Beaudouin-Laf	{VIGO}: instrumental interaction in multi-surface environme			
	2	2	Kervinen, Antti and Maunumaa, Mika and Katara, Mika	Model-Based Testing Through a {GUI}			
	2	3	Hackner, Daniel R. and Memon, Atif M.	Test case generator for {GUITAR}			
	2	4	Grechanik, Mark and Xie, Qing and Fu, Chen	Maintaining and evolving {GUI}-directed test scripts			
	2	5	Graham, Dorothy and Fewster, Mark	Experiences of Test Automation: Case Studies of Software			
		6	Gamma, Erich and Helm, Richard and Johnson, Ralp	Design patterns: elements of reusable object-oriented soft			
		7	El-Far, Ibrahim K. and Whittaker, James A.	Model-based software testing			

Incorrect

Feedback

Status: Searching for duplicates...

х

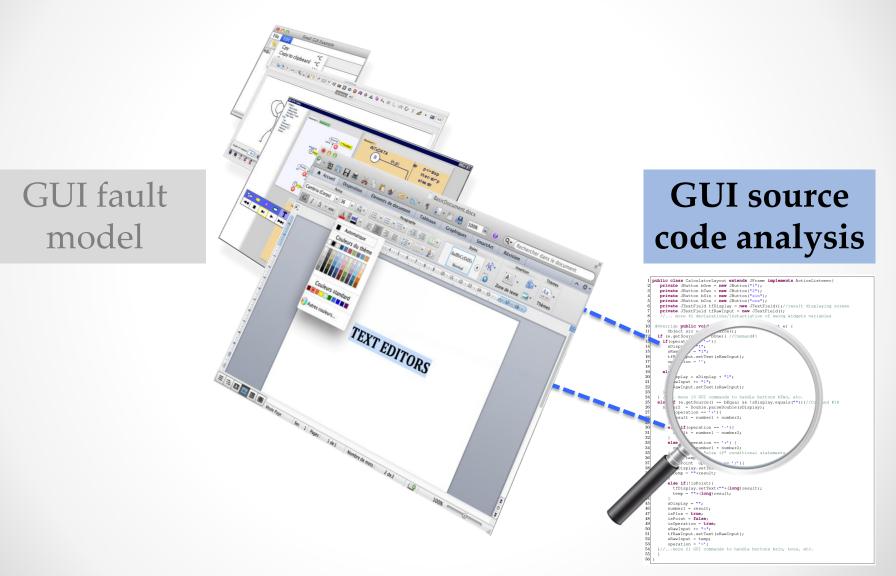
RQ2: Are GUI testing tools able to detect the classified failures?

- ✓ Most of GUI faults concern standard widgets
- X Faults that concern the interactive features such as feedback and reversibility

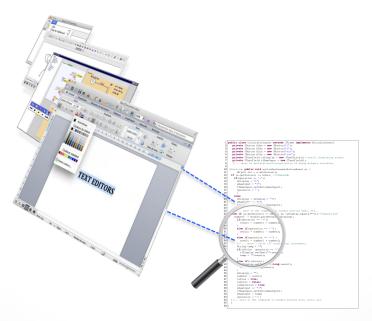
Conclusion

- An empirical study of real GUI failures
 - 279 GUI-related bug reports
- Evaluation of two GUI testing tools against
 - Real GUI failures into standard and *ad hoc* widgets
 - 65 GUI mutants derived from our fault model
 - 43 GUI mutants were not killed
- A precise analysis of standard GUI testing frameworks
 - Why GUI failures that stem from GUI faults described in our fault model were not detected?

Contributions



GUI Design Smells: The case of Blob Listener



GUI code assurance quality

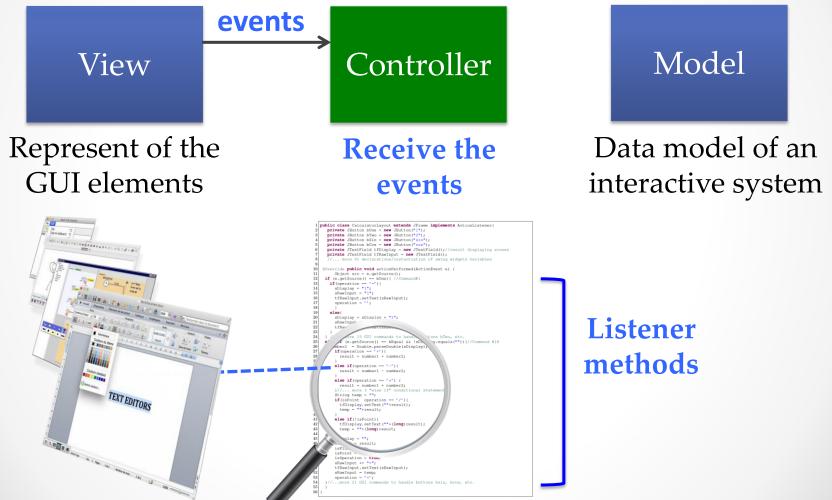
Objectives



- Identify and characterize design smells that degrade the GUI code quality
- Develop a novel static analysis to detect GUI design smells

GUI implementations





GUI implementations

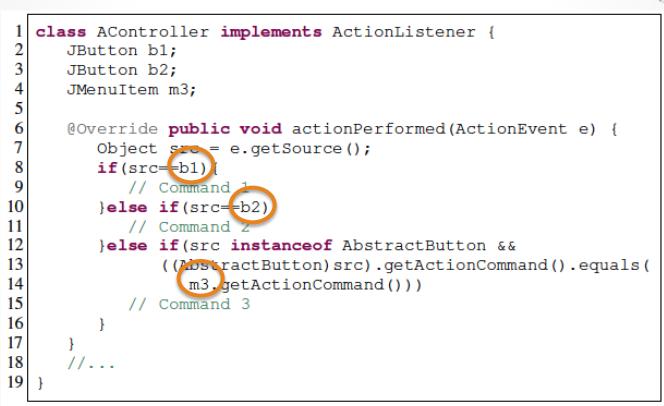
- Specific architectural design patterns
 - Organize the GUI components
 - Describe how the components interact with each other

Mode-View*

- Model-View Controller (MVC)
- Mode-View Presenter (MVP)
- Model-View-ViewModel (MVVM), etc.



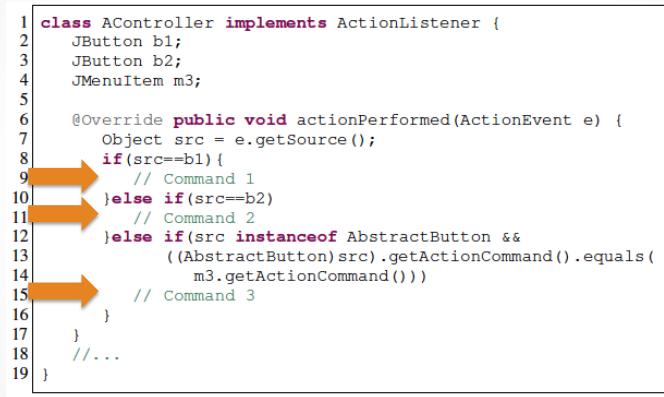
Java GUI controller



 AController manages events produced by three widgets (b1, b2, and m3) GUI source code

Blob listener

• A GUI listener produces **several** GUI commands



GUI command is a set of statements executed in reaction of a user interaction

Empirical Study on GUI listeners

13 open-source software systems

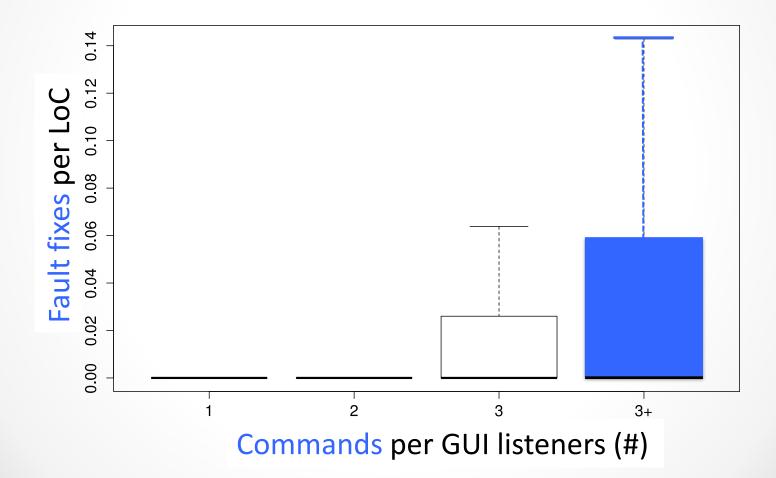
- Github repository that use an issue-tracking system
- Large Java systems
- GUI size: 858 GUI listeners

Metrics

- Average commits
- Average fault fixes
- Number of commands

Results

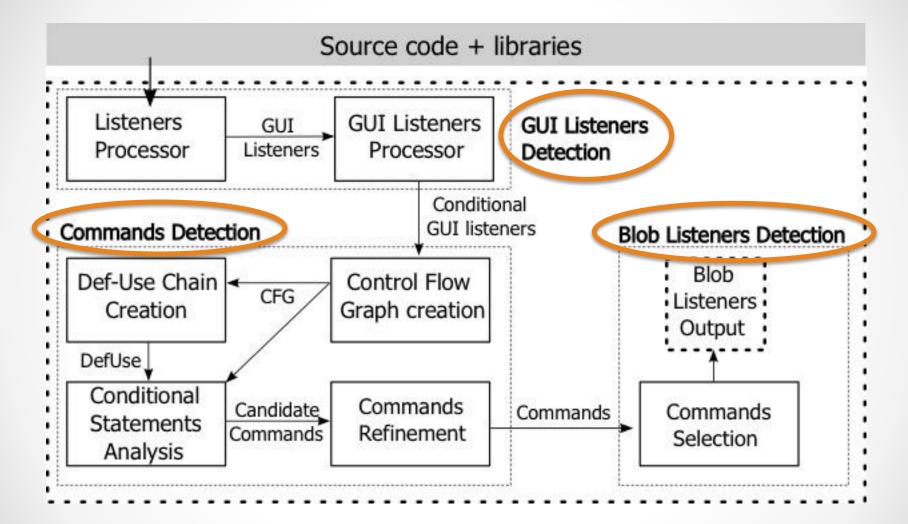
✓ The number of commands per GUI listeners has a negative impact on fault-proneness of listeners code

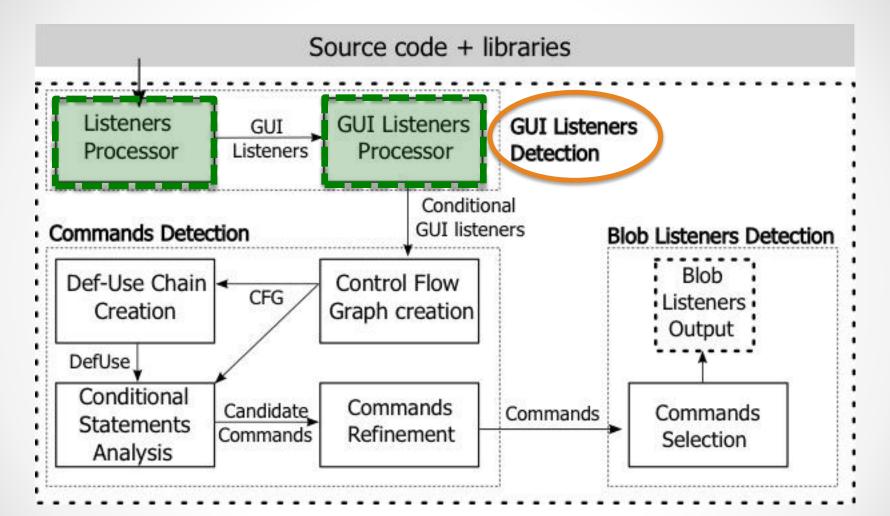


Results

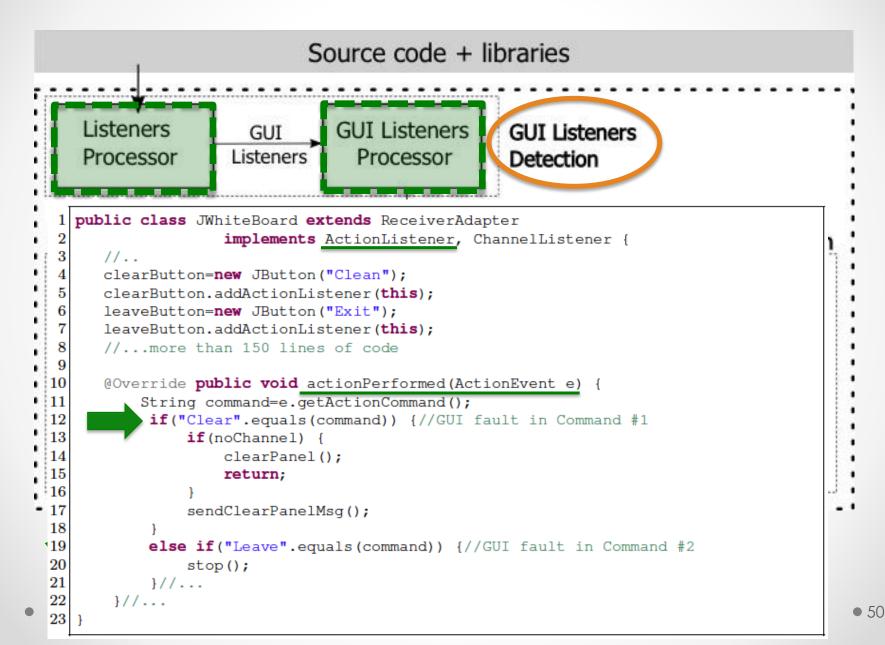
- Establish a threshold value to at least three commands per listener
 - ✓ 21% of the analyzed GUI listeners are *Blob listeners*

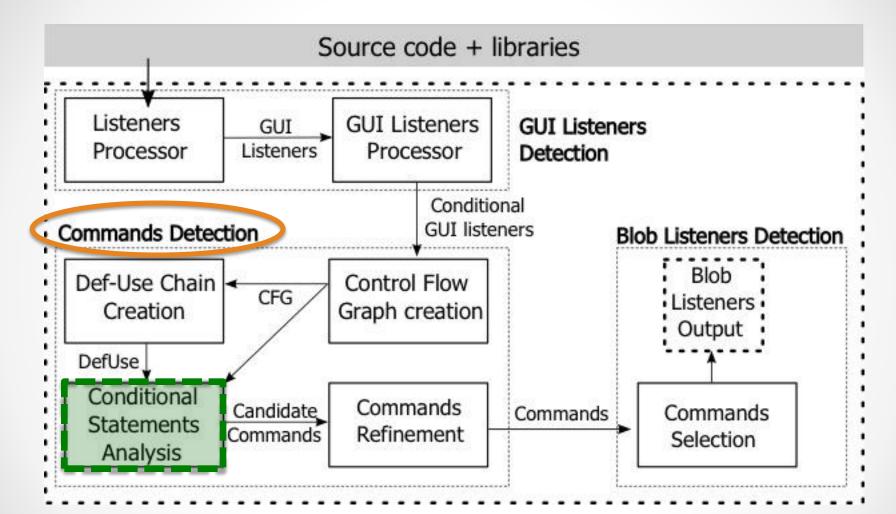
Blob Listener is a GUI listener that produces **more than two** GUI commands



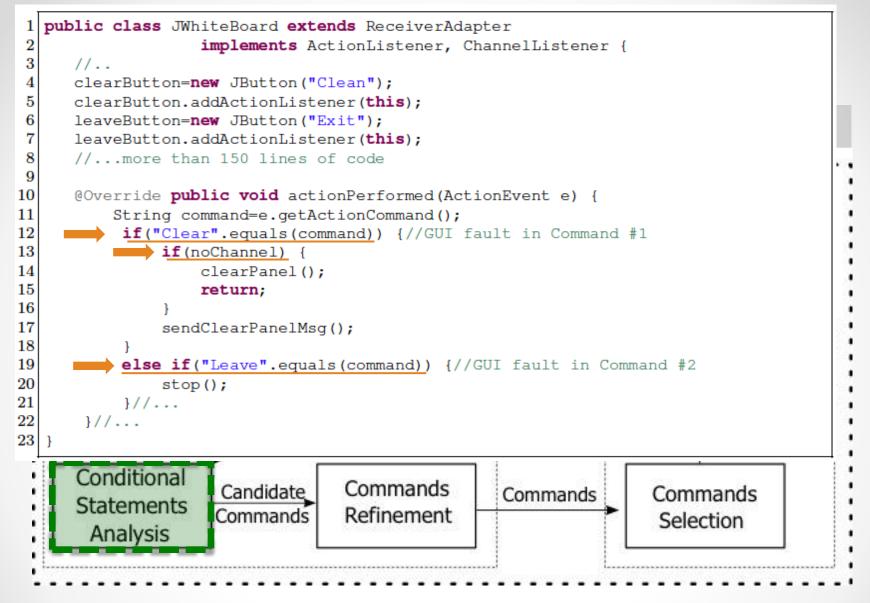


✓ GUI listeners are analyzed to identify GUI listeners that have at least one conditional statement

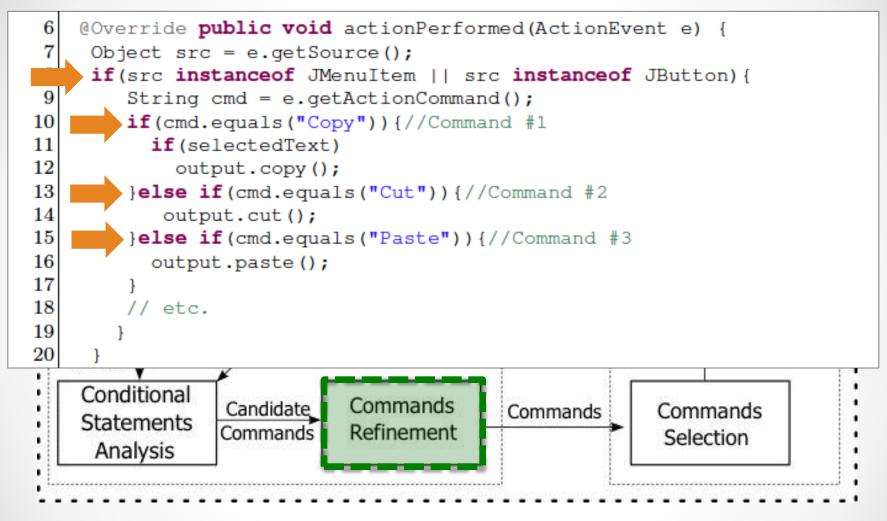




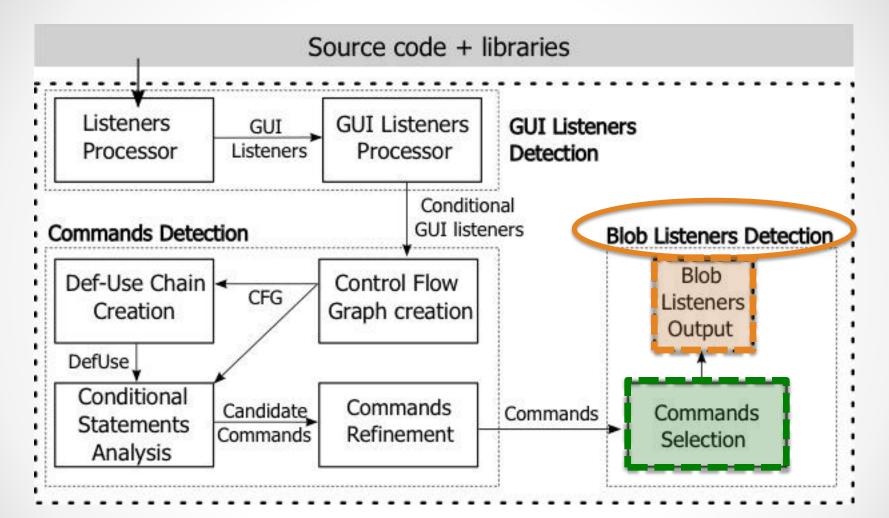
✓ The conditionals are analyzed to detect any reference to a GUI event or widget



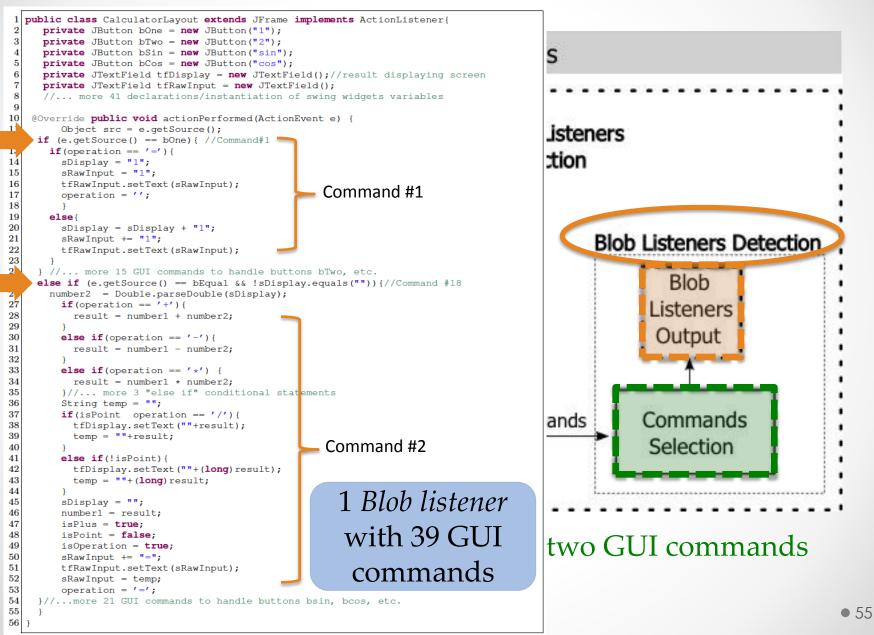
✓ The conditionals are analyzed to detect any reference to a GUI event or widget



✓ The nested commands are removed



✓ GUI listeners that contain more than two GUI commands are marked as Blob Listener



InspectorGuidget⁷

 Open-source tool as an Eclipse plug-in dedicated to Java GUI systems

Software System	Successfully Detected Blob listeners (#)	FN (#)	FP (#)	Recall (%)	Precision (%)
FastPhotoTagger	3	0	0	100.00	100.00
GanttProject	2	0	0	100.00	100.00
JaxoDraw	7	0	1	100.00	87.50
Jmol	11	1	0	91.67	81.82
TerPaint	3	0	0	100.00	100.00
TripleA	11	0	0	100.00	100 00
Overall	37	1	1	97.59	97.37

✓ 37 out of 38 *Blob listeners* were detected

⁷https://github.com/diverse-project/InspectorGuidget

Conclusion

- A new type of GUI design smell
 - Blob listener has an negative impact on fault-proneness of GUI listeners
- A novel static analysis approach
 - InspectorGuidget dedicated to Java systems
 - 37 out of 38 instances of *Blob listeners* on six real-world GUI systems
- Good coding practices to avoid the presence of *Blob listeners*

Conclusions & Perspectives

Conclusions

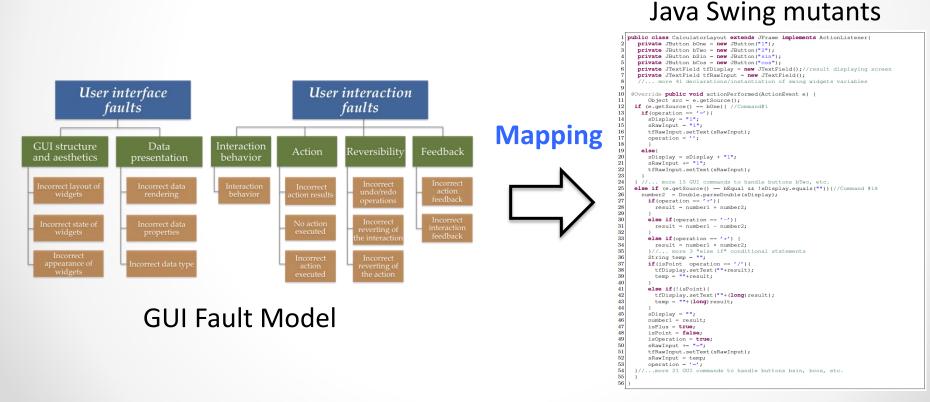
- GUI fault model
 - 279 GUI-related bug reports of five interactive open-source systems
 - Evaluation of GUI testing frameworks against real GUI failures and GUI mutants
- An automatic detection of a new type of GUI design smell
 - Blob listener that degrades the GUI code quality
 - InspectorGuidget detected 37 out 38 instances of *Blob listeners*

Conclusions

- Experiment and tools
 - A complete data set
 - GUI systems that have several interactive features
 - Empirial studies of GUI implementations

Perspectives

- Domain-specific mutants
 - Mapping between GUI faults and specific GUI toolkits



Perspectives

GUI design smells

 A set of checking rules to check automatically for potential defects in GUI code



Perspectives

GUI design smells



Bug finders Findbugs PMD, etc.

Publications

- Valéria Lelli, Arnaud Blouin, and Baudry Benoit. Classifying and qualifying GUI defects. In Software Testing, Verification and Validation (ICST), 2015 IEEE Eighth International Conference, pages 1–10, April 2015.
- Valéria Lelli, Arnaud Blouin, Baudry Benoit, and Fabien Coulon. On modelbased testing advanced GUIs. In 11th Workshop on Advances in Model Based Testing (A-MOST), pages 1–10, April 2015. *Best paper award*.
- Valéria Lelli. Challenges of testing for critical interactive systems. In Software Testing, Verification and Validation (ICST), 2013 IEEE Sixth International Conference. Doctoral Symposium, pages 509–510, March 2013.
- Valéria Lelli, Arnaud Blouin, Baudry Benoit, Fabien Coulon, and Olivier Beaudoux. Automatic Detection of GUI Design Smells: The Case of Blob Listener. Submitted to Software Testing, Verification and Validation Conference (ICST) 2016.