

Impact of the Visitor Pattern on Program Comprehension and Maintenance

Sébastien Jeanmart



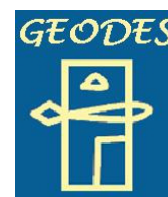
Univ. de Namur

Yann-Gaël Guéhéneuc



DGIGL, École
Polytech. de Montréal

Houari Sahraoui



DIRO
Univ. de Montréal

Naji Habra

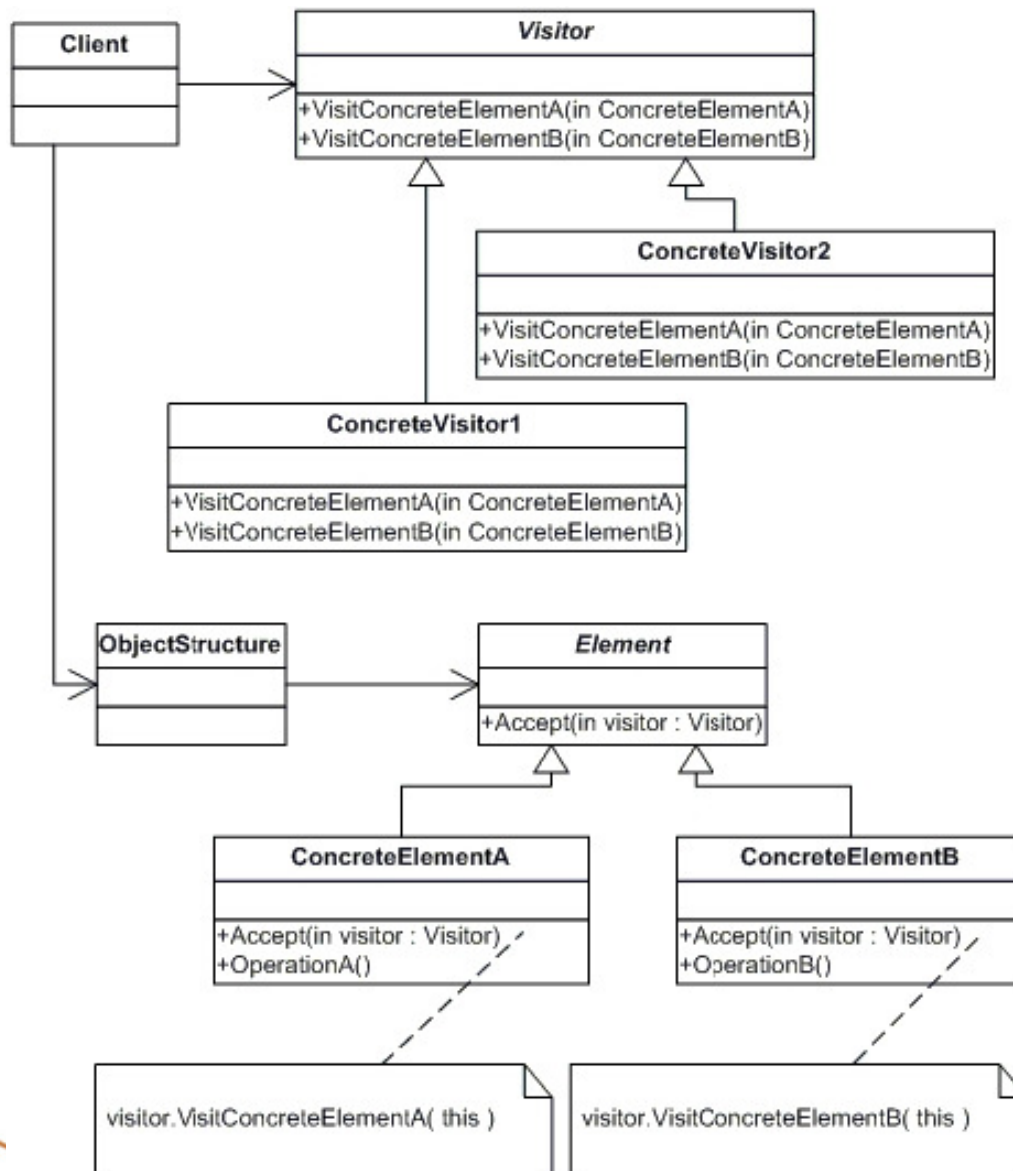


Univ. de Namur

Motivation

- Much work based on the premises that patterns improve
 - analyzability
 - changeability
- Some studies suggested that patterns may impact negatively
 - development and maintenance activities
- Goal of this study : Impact of Visitor pattern
 - comprehension
 - maintenance

Visitor Pattern



Study Objectives

- Study the impact of Visitor pattern
 - at the design level (UML class diagrams)
 - on comprehension and modification tasks
- From two perspectives
 - Presence vs. absence of the pattern
 - The layout of the classes involved in the pattern

Hypotheses

- Comprehension
 - HCa1
 - A class diagram with the Visitor reduces the subjects' effort during program comprehension when compared to one without it
 - HCa2
 - A class diagram using the canonical representation of the Visitor reduces the subjects' effort during program comprehension when compared to one using the Visitor with another layout

Hypotheses

- Maintenance

- HMa1

- A class diagram with the Visitor reduces the subjects' effort during program modification when compared to a class diagram without it

- HMa2

- A class diagram using the canonical representation of the Visitor reduces the subjects' effort during program modification when compared to one using the Visitor with another layout

Variables

- Independent variables
 - Design Alternative
 - Visitor and canonical representation (CP)
 - Visitor and random layout (MP)
 - without Visitor (NP)
 - Tasks
 - Comprehension (C)
 - Modification (M)
- Mitigating variables
 - UML Knowledge: (1) fair/good (2) very good
 - DP Knowledge: (1) fair/good (2) very good

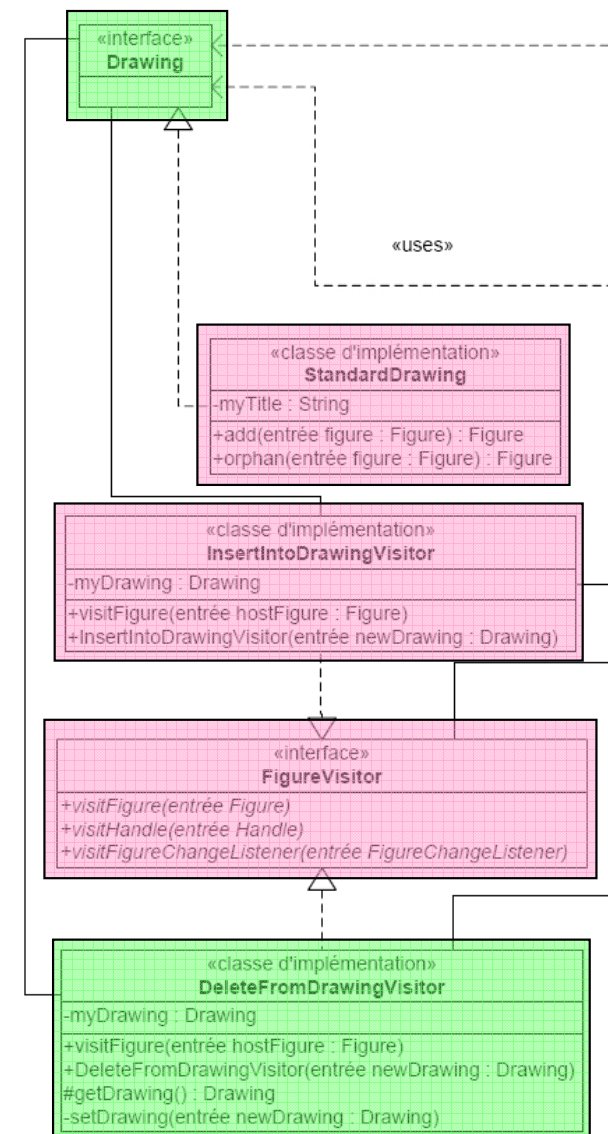
Variables

- **Dependent variables**
 - Eye tracking system (EyeLink II from SR Research)
 - Two cameras on a headband
 - Detection of eye-movements using infra-red ray reflection on the subject's cornea
 - Sensors help translating eye-movements into screen positions
 - Raw data
 - Fixations: stabilizations of the eye during a gaze, characterized by a duration
 - Saccades: movements from one fixation to another
 - Subject's effort
 - Derived from the number and durations of fixations, and saccades



Variables

- **Dependent variables**
 - From screen coordinate to diagrams
 - Fixations and saccades grouped by areas of interest corresponding to class representations in a diagram
 - Two categories of classes (areas)
 - Relevant: necessary to perform the task
 - Non-relevant: otherwise



Variables

- **Dependent variables**
 - ADRF and ADNRF (Average Duration of Relevant (resp. Non-Relevant) Fixations)
 - **Conjecture**
 - First phase: subjects look at the different classes of the diagram to identify relevant ones
 - Second phase: they use mainly relevant classes to perform the task
 - By hypothesis, Visitor reduces the duration of the second phase

- **Calculation**

$$ADRF = \frac{\sum_{c \in \{Rel. Classes\}} d(c)}{\#\{Rel. Classes\}}$$

$$ADNRF = \frac{\sum_{c \in \{NonRel. Classes\}} d(c)}{\#\{NonRel. Classes\}}$$

$d(c)$: total duration of the fixations made to a class c

Variables

- Dependent variables

- NRRF (Normalized Rate of Relevant Fixations)

- Conjecture

- When subject look often to the relevant classes, it means that he is making a lot of the effort to perform the task
- The number of relevant fixations is normalized by the size of the diagram
- By hypothesis, Visitor reduces the relative number of relevant fixations

- Calculation

$$NRRF = \frac{\frac{\sum_{c \in \{Rel. Classes\}} f(c)}{\#\{Rel. Classes\}}}{\frac{\sum_{c \in \{Rel. Classes\} \cup \{NonRel. Classes\}} f(c)}{\#\{Rel. Classes\} + \#\{NonRel. Classes\}}}$$

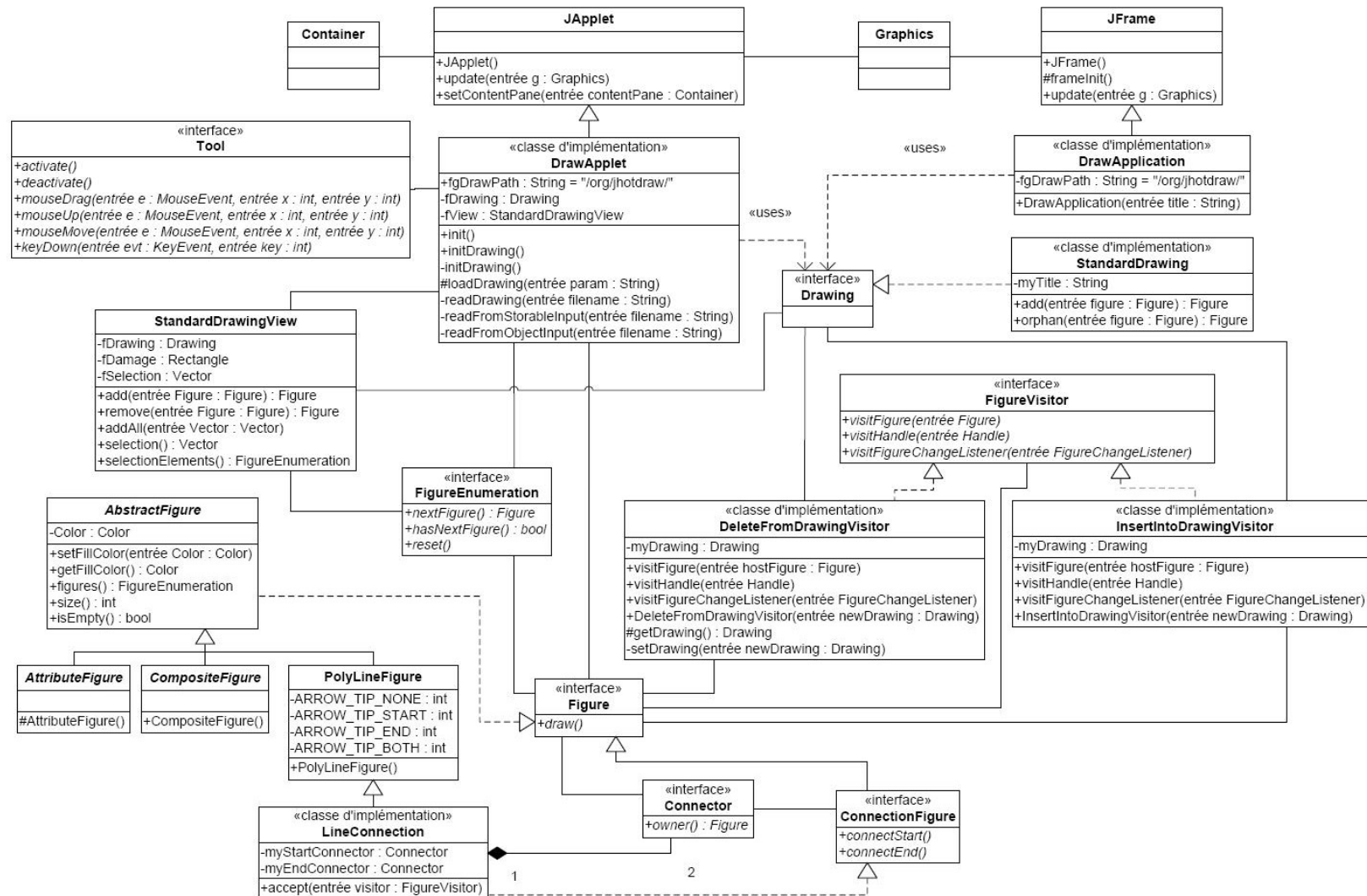
$f(c)$: number of the fixations made to a class c

Objects and Subjects

- **Objects**
 - Reverse-engineered class diagrams
 - Three open-source programs/libraries
 - JHotDraw, JRefactory, and PADL
- **Subjects**
 - 24 students (graduate) in three groups
 - Experience with software architecture and UML diagrams

	No Pattern V_n	ClassicalPattern V_p	ModifiedPattern V_m
JHotDraw	G_1	G_2	G_3
JRefactory	G_3	G_1	G_2
PADL	G_2	G_3	G_1

Tasks



Is it possible to change the color of a LineConnection? Why ?

Study Results

- Comprehension
 - The presence of Visitor, with both layouts, does not change significantly the effort for comprehension tasks

(ms)	CP_C	MP_C	NP_C	$\Delta 1$ (NP_C, CP_C)	$\Delta 2$ (NP_C, MP_C)	sig. ($\Delta 1$)	sig. ($\Delta 2$)
ADRF	12,766	12,252	13,332	4%	8%	0.807	0.592
ADNRF	3,355	3,578	3,388	1%	-6%	0.711	0.635
(%)	CP_C	MP_C	NP_C	p-value (HC01)		p-value (HC02)	
NRRF	75.77	77.95	81.49	0.221		0.663	

Study Results

- Maintenance
 - The presence of Visitor (only with the canonical layout) does change significantly the effort for maintenance tasks

(ms)	<i>CP_M</i>	<i>MP_M</i>	<i>NP_M</i>	$\Delta 1$ (<i>NP_M, CP_M</i>)	$\Delta 2$ (<i>NP_M, MP_M</i>)	sig. ($\Delta 1$)	sig. ($\Delta 2$)
ADRF	12,138	13,175	18,724	35%	30%	0.045	0.076
ADNRF	4,804	4,024	4,421	-9%	9%	0.672	0.626
(%)	<i>CP M</i>	<i>MP M</i>	<i>NP M</i>	<i>p</i> -value (<i>HM01</i>)		<i>p</i> -value (<i>HM02</i>)	
NRRF	71.97	73.46	80.18	0.027		0.725	

Study Results

- Combined effect
 - UML knowledge
 - No significant impact (alone or combined)
 - DP knowledge
 - No significant impact (alone or combined)

Threats to the Validity

- Internal validity
 - Tasks given in random and different orders to subjects (learning and fatigue effects)
 - Limited time for the experiment (fatigue effect)
 - Analysis of eye-movement movies to detect cases of decalibration (instrumentation)
 - No communication (about the study) between the subjects during the study period (treatment learning)

Threats to the Validity

- Construct validity
 - Diagrams extracted from three programs with different application domains (mono-operation bias)
 - Two dependent variables: ADRF-ADNRF and NRRF(mono-method bias)
 - Subjects not informed before hand about the study goal (hypothesis guessing)
 - Subjects reassured about the absence of risks and anonymous identification (apprehension)

Threats to the Validity

- External validity
 - Graduate students have knowledge of UML and design patterns comparable to most junior professionals (interaction between selection and treatment)
 - Actual UML diagrams with a usual size of 20 classes ((interaction between setting and treatment)

Threats to the Validity

- Conclusion validity
 - Quiet laboratory and pre-study trials (random irrelevancy in setting)
 - 24 observations per design alternative
 - Heterogeneous subjects with different 2 levels of UML and DP knowledge, different GPAs, and with or without working experience
 - Combined effect analysis

Conclusion

- Analysis of the impact of Visitor for comprehension and modification tasks
- Considering the layout
- Using an eye-tracking system (objective effort measurement)
- Results
 - Visitor does not reduce the comprehension effort
 - Visitor does reduce the modification effort when presented in the canonical layout

Future Work

- Replications
 - Other variables
 - Other patterns
 - Professional subjects