


Analyzing the inconsistency between comments and source code in Pharo

Bachelor Thesis by Sabine Brunner
Supervised by Pooja Rani

A dark blue diagonal gradient bar that starts from the bottom left corner and extends towards the top right corner, covering the lower half of the slide.

Problems with comments

- Comments don't get written because...
 - Time constraints
 - Laziness
 - Pride (e.g. "Good code speaks for itself")
- Comments are wrong because...
 - Software evolution
 - Refactoring tools

Problems with comments

“Type a class comment for LOCell and accept it; you can always improve it later.”

(<http://pharo.gforge.inria.fr/PBE1/PBE1ch3.html>, 27.10.2018)

“I wish there was some trick to make other necessary and productive tasks feel as good as programming.”

- *Reddit user Deto*

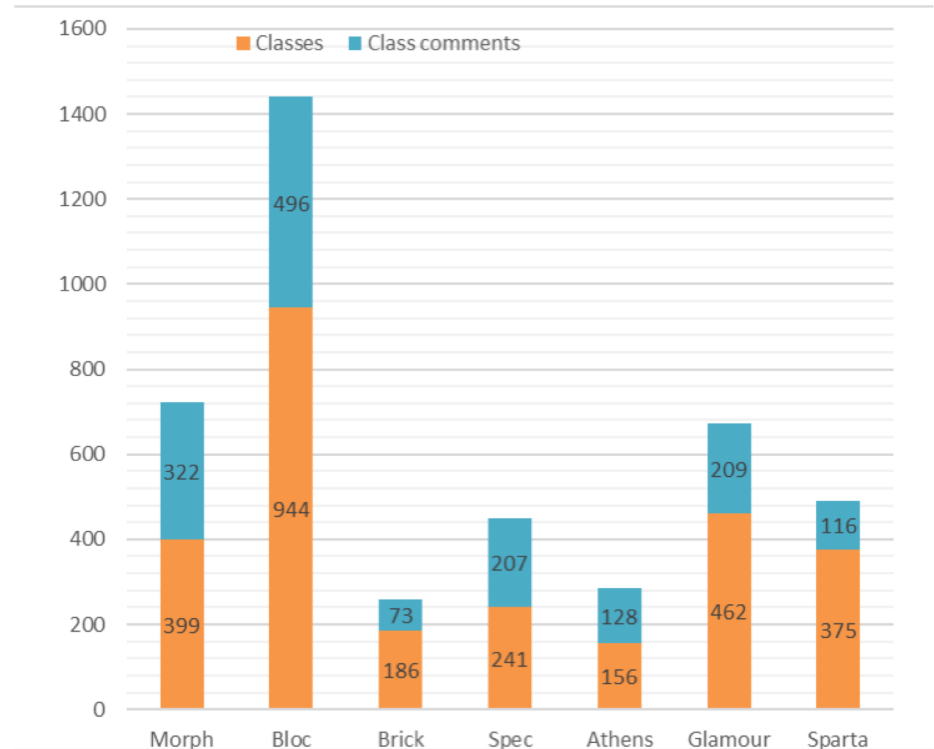
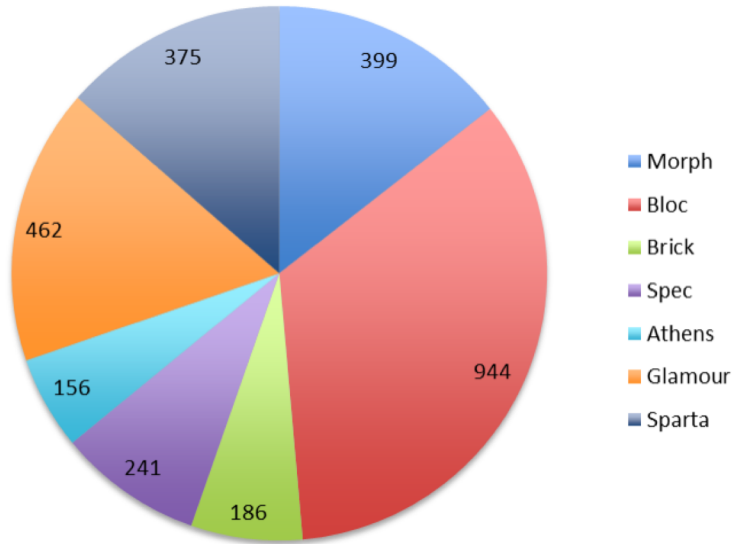
(https://www.reddit.com/r/programming/comments/1ff643/why_do_programmers_hate_documenting/, 27.10.2018)

“Captain obvious tells this is an array of CairoGlyph(s), corresponds to cairo_glyph_t structure”

- *CairoGlyphsArray (Athens-Cairo)*

Dataset

User Interface libraries in Pharo



Tools and Techniques

Stanford CoreNLP Server

- Part-of-Speech tagging
- Constituency parsing
- Dependency parsing

<https://stanfordnlp.github.io/CoreNLP/corenlp-server.html>

<http://corenlp.run/>

The Pharo Class comment template

Please comment me using the following template inspired by Class Responsibility Collaborator (CRC) design:

For the Class part: State a one line summary. For example, "I represent a paragraph of text".

For the Responsibility part: Three sentences about my main responsibilities - what I do, what I know.

For the Collaborators Part: State my main collaborators and one line about how I interact with them.

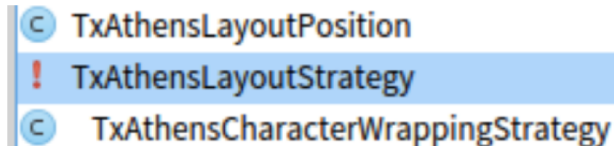
Public API and Key Messages

- message one
- message two
- (for bonus points) how to create instances.

One simple example is simply gorgeous.

Internal Representation and Key Implementation Points.

- Class Responsibility Collaborator (CRC) Style
- First person narrator “I” explains
 - Responsibilities
 - Collaborators
 - Key messages



● TxAthensLayoutPosition
! TxAthensLayoutStrategy
● TxAthensCharacterWrappingStrategy

Manual Analysis examples

"Athens text background"

– *AthensTextBackground* (Athens-Text)

"do not use this class. this is subject of change or
removal"

- *AthensCairoText (Athens-Cairo)*

"corresponds to cairo_glyph_t structure:

```
typedef struct {  
    unsigned long    index;  
    double           x;  
    double           y;  
} cairo_glyph_t;
```

- CairoGlyphsArray (Athens-Cairo)

"Before start rendering, **you must set:** - canvas, font
[...]"

- *CairoFreetypeFontRenderer (Athens-Cairo)*

"Tools for debugging:

Display the images momentarily under program control (for positioning) (self is an instance).

self state: #on. self state: #off.

self state: #pressed. self state: #off."

- *ThreePhaseButtonMorph (Athens-Morphic)*

"**and** [...] **Things get messed up** such that the last return character seems to be gone. In this state, if you position the cursor immediately after the 2, and press the right arrow, the cursor jumps to the beginning of line 2... oof. (report by Doug Way) [...] **we do need to clean this baby up some day. - Bob**"

- Paragraph (Athens-Morphic)

Template vs. actual user usage

Template:

- Responsibilities
- Collaborators
- Key messages

Observed usage:

- Warnings
- Bug reports
- Discourse
- Online documentation
- Instructions

Research Questions

- Which parts of the template are used?
- What are the inconsistencies between the class comments and the source code?
- What is the overall quality of the class comments?
 - Do they make use of standard domain terms?
 - Do they mention referenced classes?
 - Do they explain their main messages?

How can we make good commenting easier?

Next Steps

- Analysis of References:
 - which classes/methods are referenced in a comment?
 - how are they referenced? What is said about them?
- Analysis of the Structure:
 - to which extent do the comments make use of the template?
 - do most users follow the same structure?
- Analysis of content/intent:
 - What is the intent of class comments?
 - Which information is mentioned often, which isn't?

Discussion

A dark blue diagonal gradient bar that starts from the bottom left corner and extends towards the top right corner, covering the lower half of the slide.