A Framework for Estimating Interest on Technical Debt by Monitoring Developer Activity Related to Code Comprehension

Vallary Singh, University Of Delaware
Will Snipes, ABB Corporate Research
Nicholas A. Kraft, ABB Corporate Research
Technical Debt

Cost to repay debt

Extra Cost while working around the debt

Our Focus: Estimating Interest
Current Approaches

• Principal Only:
  – Debt related to architecture, code structure, code complexity and code smells (Kruchten et. al. 2012)

• Principal and Interest:
  – Maintainability ratings (Nugorho et. al. 2011)
  – Extra effort to modify unnecessarily complex code (Curtis et. al. 2012)

➢ Little focus on estimating interest.
➢ Current estimates miss impact of debt on developers’ code comprehension **effort.**
Hypothesis

While static code metrics have been shown useful in quantifying Principal for estimating Technical Debt,

Our Hypothesis:

_Comprehension effort metrics are useful for quantifying Interest._
Contributions

Framework to compute Comprehension Effort using Developer Activity Logs:

– Define Class-based Comprehension Effort Metrics computed from developer logs

– Perform initial case study
  • 9 months of developer logs
  • 9888 classes in project
Approach

Goal: Relate static code metrics to developer activity

- Monitoring Data
- Establish Sessions
- Define Class Based Metrics
- Identify Central Class(es) that were changed
- Relate to Static Metrics of Central Class

- #Sessions
- #Visits to central class
- #Other class visits
- Time spent in central class
- Time spent in other classes
Monitoring IDE Activity with Blaze*

<table>
<thead>
<tr>
<th>Time-stamp</th>
<th>User</th>
<th>Event</th>
<th>Artifact</th>
</tr>
</thead>
<tbody>
<tr>
<td>22:04:51</td>
<td>N3</td>
<td>View.SourceFile</td>
<td>1acc7366.cs/10</td>
</tr>
<tr>
<td>22:04:52</td>
<td>N3</td>
<td>View.OnChangeCaretLine</td>
<td>1acc7366.cs/14</td>
</tr>
<tr>
<td>22:04:53</td>
<td>N3</td>
<td>View.OnChangeCaretLine</td>
<td>1acc7366.cs/16</td>
</tr>
<tr>
<td>22:04:58</td>
<td>N3</td>
<td>Menu.ViewCallHierarchy</td>
<td>1acc7366.cs/16</td>
</tr>
<tr>
<td>22:05:00</td>
<td>N3</td>
<td>View.OnChangeCaretLine</td>
<td>1acc7366.cs/20</td>
</tr>
<tr>
<td>22:05:19</td>
<td>N3</td>
<td>View.SourceFile</td>
<td>81c2db1a.cs/1</td>
</tr>
<tr>
<td>22:05:22</td>
<td>N3</td>
<td>Edit.Find</td>
<td>81c2db1a.cs/1</td>
</tr>
<tr>
<td>22:05:30</td>
<td>N3</td>
<td>Edit.FindNext</td>
<td>81c2db1a.cs/20</td>
</tr>
</tbody>
</table>

* Developed at ABB Inc (Snipes et.al. 2014)
Establish Sessions

Class accesses as time progresses

1st Session for A

2nd Session for A
Defining Central Class Based on Effort Metrics

<table>
<thead>
<tr>
<th>Class</th>
<th>#Sessions</th>
<th>#Class Visits</th>
<th>#Other Class Accesses</th>
<th>Time Spent in Class</th>
<th>Time Spent in Other Classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>74</td>
<td>294</td>
<td>78</td>
<td>39 hours</td>
<td>17 hours</td>
</tr>
<tr>
<td>B</td>
<td>32</td>
<td>88</td>
<td>52</td>
<td>3 hours</td>
<td>9 hours</td>
</tr>
<tr>
<td>C</td>
<td>28</td>
<td>77</td>
<td>52</td>
<td>4.5 hours</td>
<td>9 hours</td>
</tr>
</tbody>
</table>

Comprehension Effort Metrics
### Defining Central Class Based on Effort Metrics

#### Comprehension Effort Metrics

Class A visited often and over long period of time

<table>
<thead>
<tr>
<th>Class</th>
<th>#Sessions</th>
<th>#Class Visits</th>
<th>#Other Class Accesses</th>
<th>Time Spent in Class</th>
<th>Time Spent in Other Classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>74</td>
<td>294</td>
<td>78</td>
<td>39 hours</td>
<td>17 hours</td>
</tr>
<tr>
<td>B</td>
<td>32</td>
<td>88</td>
<td>52</td>
<td>3 hours</td>
<td>9 hours</td>
</tr>
<tr>
<td>C</td>
<td>28</td>
<td>77</td>
<td>52</td>
<td>4.5 hours</td>
<td>9 hours</td>
</tr>
</tbody>
</table>
**Defining Central Class Based on Effort Metrics**

Even though No of sessions is low No of Class visits is comparatively high

<table>
<thead>
<tr>
<th>Class</th>
<th>#Sessions</th>
<th>#Class Visits</th>
<th>#Other Class Accesses</th>
<th>Time Spent in Class</th>
<th>Time Spent in Other Classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>74</td>
<td>294</td>
<td>78</td>
<td>39 hours</td>
<td>17 hours</td>
</tr>
<tr>
<td>B</td>
<td>32</td>
<td>88</td>
<td>52</td>
<td>3 hours</td>
<td>9 hours</td>
</tr>
<tr>
<td>C</td>
<td>28</td>
<td>77</td>
<td>52</td>
<td>4.5 hours</td>
<td>9 hours</td>
</tr>
</tbody>
</table>

**Comprehension Effort Metrics**
### Defining Central Class Based on Effort Metrics

Can be interpreted as comprehension time

<table>
<thead>
<tr>
<th>Class</th>
<th>#Sessions</th>
<th>#Class Visits</th>
<th>#Other Class Accesses</th>
<th>Time Spent in Class</th>
<th>Time Spent in Other Classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>74</td>
<td>294</td>
<td>78</td>
<td>39 hours</td>
<td>17 hours</td>
</tr>
<tr>
<td>B</td>
<td>32</td>
<td>88</td>
<td>52</td>
<td>3 hours</td>
<td>9 hours</td>
</tr>
<tr>
<td>C</td>
<td>28</td>
<td>77</td>
<td>52</td>
<td>4.5 hours</td>
<td>9 hours</td>
</tr>
</tbody>
</table>

### Comprehension Effort Metrics
## Relating Effort to Static Metrics

<table>
<thead>
<tr>
<th>Class</th>
<th>#Class Base</th>
<th>#Class Coupled</th>
<th>#Class Derived</th>
<th>#LOC (Top Rank %)</th>
<th>#Declared Method (Top Rank %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>10</td>
<td>272</td>
<td>1</td>
<td>0.05%</td>
<td>0.13%</td>
</tr>
<tr>
<td>B</td>
<td>3</td>
<td>98</td>
<td>0</td>
<td>0.12%</td>
<td>2.20%</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>32</td>
<td>0</td>
<td>0.70%</td>
<td>6.90%</td>
</tr>
</tbody>
</table>

**Class Level Code Structure Metrics**
# Relating Effort to Static Metrics

## Class Level Code Structure Metrics

<table>
<thead>
<tr>
<th>Class</th>
<th>#Class Base</th>
<th>#Class Coupled</th>
<th>#Class Derived</th>
<th>#LOC (Top Rank %)</th>
<th>#Declared Method (Top Rank %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>10</td>
<td>272</td>
<td>1</td>
<td>0.05%</td>
<td>0.13%</td>
</tr>
<tr>
<td>B</td>
<td>3</td>
<td>98</td>
<td>0</td>
<td>0.12%</td>
<td>2.20%</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>32</td>
<td>0</td>
<td>0.70%</td>
<td>6.90%</td>
</tr>
</tbody>
</table>

**Code Smell**

---

*University of Delaware*
### Relating Effort to Static Metrics

<table>
<thead>
<tr>
<th>Class</th>
<th>#Class Base</th>
<th>#Class Coupled</th>
<th>#Class Derived</th>
<th>#LOC (Top Rank %)</th>
<th>#Declared Method (Top Rank %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>10</td>
<td>272</td>
<td>1</td>
<td>0.05%</td>
<td>0.13%</td>
</tr>
<tr>
<td>B</td>
<td>3</td>
<td>98</td>
<td>0</td>
<td>0.12%</td>
<td>2.20%</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>32</td>
<td>0</td>
<td>0.70%</td>
<td>6.90%</td>
</tr>
</tbody>
</table>

### Class Level Code Structure Metrics

<table>
<thead>
<tr>
<th>Class</th>
<th>#Sessions</th>
<th>#Class Visits</th>
<th>#Other Class Accesses</th>
<th>Time Spent in Class</th>
<th>Time Spent in Other Classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>74</td>
<td>294</td>
<td>78</td>
<td>39 hours</td>
<td>17 hours</td>
</tr>
<tr>
<td>B</td>
<td>32</td>
<td>88</td>
<td>52</td>
<td>3 hours</td>
<td>9 hours</td>
</tr>
<tr>
<td>C</td>
<td>28</td>
<td>77</td>
<td>52</td>
<td>4.5 hours</td>
<td>9 hours</td>
</tr>
</tbody>
</table>

### Comprehension Effort Metrics
## Relating Effort to Static Metrics

### Class Level Code Structure Metrics

<table>
<thead>
<tr>
<th>Class</th>
<th>#Sessions</th>
<th>#Class Visits</th>
<th>#Other Class Accesses</th>
<th>Time Spent in Class</th>
<th>Time Spent in Other Classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>74</td>
<td>294</td>
<td>78</td>
<td>39 hours</td>
<td>17 hours</td>
</tr>
<tr>
<td>B</td>
<td>32</td>
<td>88</td>
<td>52</td>
<td>3 hours</td>
<td>9 hours</td>
</tr>
<tr>
<td>C</td>
<td>28</td>
<td>77</td>
<td>52</td>
<td>4.5 hours</td>
<td>9 hours</td>
</tr>
</tbody>
</table>

### Comprehension Effort Metrics

<table>
<thead>
<tr>
<th>Class</th>
<th>#Class Base</th>
<th>#Class Coupled</th>
<th>#Class Derived</th>
<th>#LOC (Top Rank %)</th>
<th>#Declared Method (Top Rank %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>10</td>
<td>272</td>
<td>1</td>
<td>0.05%</td>
<td>0.13%</td>
</tr>
<tr>
<td>B</td>
<td>3</td>
<td>98</td>
<td>0</td>
<td>0.12%</td>
<td>2.20%</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>32</td>
<td>0</td>
<td>0.70%</td>
<td>6.90%</td>
</tr>
</tbody>
</table>

### Highly Coupled Class

- A
- B
- C

---

**Note:** The metrics are derived from static code analysis and class level code structure, and they provide insights into the effort required for comprehension and maintenance.
Relating Effort to Static Metrics

### Class Level Code Structure Metrics

<table>
<thead>
<tr>
<th>Class</th>
<th>#Class Base</th>
<th>#Class Coupled</th>
<th>#Class Derived</th>
<th>#LOC (Top Rank %)</th>
<th>#Declared Method (Top Rank %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>10</td>
<td>272</td>
<td>1</td>
<td>0.05%</td>
<td>0.13%</td>
</tr>
<tr>
<td>B</td>
<td>3</td>
<td>98</td>
<td>0</td>
<td>0.12%</td>
<td>2.20%</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>32</td>
<td>0</td>
<td>0.70%</td>
<td>6.90%</td>
</tr>
</tbody>
</table>

### Comparable Value

### Comprehension Effort Metrics

<table>
<thead>
<tr>
<th>Class</th>
<th>#Sessions</th>
<th>#Class Visits</th>
<th>#Other Class Accesses</th>
<th>Time Spent in Class</th>
<th>Time Spent in Other Classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>74</td>
<td>294</td>
<td>78</td>
<td>39 hours</td>
<td>17 hours</td>
</tr>
<tr>
<td>B</td>
<td>32</td>
<td>88</td>
<td>52</td>
<td>3 hours</td>
<td>9 hours</td>
</tr>
<tr>
<td>C</td>
<td>28</td>
<td>77</td>
<td>52</td>
<td>4.5 hours</td>
<td>9 hours</td>
</tr>
</tbody>
</table>
Future Work

• Study of correlation static vs log based comprehension
• Large scale statistical study
• Additional comprehension metrics
• Conduct validation study

Acknowledgements: Thank you ABB for supporting this work.