UML-based Object-Oriented Metrics for Architecture Complexity Analysis

Peter In\textsuperscript{1}, SangEun Kim\textsuperscript{2}, Matthew Barry\textsuperscript{3}

\textsuperscript{1} Dept. of Computer Science, Texas A&M University, College Station, TX 77840 USA, Tel: (979) 458-1547, hohin@cs.tamu.edu
\textsuperscript{2} Dept. of Computer Science, Texas A&M University, College Station, TX 77840 USA, Tel: (979) 845-5439, sangeunk@cs.tamu.edu
\textsuperscript{3} United Space Alliance, 600 Gemini, Houston, TX 77058 USA, Tel: (281) 282-3960, Matthew.R.Barry@USAHQ.UnitedSpaceAlliance.com
Today’s Agenda

- Motivations/Solution Approach
- Proposed Metrics
- Tool Support
- Experimental Results
- Conclusion
Research Issues

1. What metrics will be helpful to a project manager early in the development lifecycle?

2. How can such metrics information be collected?

3. How can the generated metrics information be utilized?
Solution Approach

- UML-based Automatic OO Metrics Counter

Input
- UML diagrams: class, use case, component, deployment, state chart, activity, sequence, collaboration and package

Output
- Key metrics: model complexity, evolution, maturity, breakage, rework, etc.
Proposed Metrics

- **OO Concepts:** Classes, Methods, Inheritance, Polymorphism
- **Metrics Tree**

![Metrics Tree Diagram]

- **Metrics Tree Nodes:**
  - PRIMITIVE
  - FAULT PRONENESS
    - CLS
    - REL
    - WMC
    - NOC
    - DIT
  - QUALITY MEASURE
    - MHF AHF
    - MIF AIF
    - OLC
  - COUPLING
    - CLC
    - PLC
  - USECASE
    - NOA NOUC NOUCA

---

*Copyright © 2003 TAMU*
Primitive Metrics

- To provide brief and basic complexity information
  - TNC (Total Number of Class) = \( \sum_{i=1}^{n} tnc_i \)
  - TNIR (Total Number of Inheritance Relationships) = \( \sum_{i=1}^{n} tnir_i \)
  - TNRR (Total Number of Realization Relationships) = \( \sum_{i=1}^{n} tnrr_i \)
  - TNUR (Total Number of Use Relationships) = \( \sum_{i=1}^{n} tnur_i \)
  - TNA (Total Number of Associations) = \( \sum_{i=1}^{n} tna_i \)
  - TNR (Total Number of Roles) = \( \sum_{i=1}^{n} tnr_i \)
  - TNO (Total Number of Operation) = \( \sum_{i=1}^{n} tno_i \)
  - TNP (Total Number of Parameters) = \( \sum_{i=1}^{n} tnp_i \)
  - TNCA (Total Number of Class Attributes) = \( \sum_{i=1}^{n} tnca_i \)
Fault-Proneness Metrics

- To predict class’s fault-proneness
  - WMC (Weighted Method per Class) = \( \sum_{i=1}^{n} c_i \)
    where, \( c_i \) is the complexity of the methods
  - NOC (Number of Children per Class) = \( \sum_{i=1}^{n} sc_i \)
    where, \( sc_i \) is the number of immediate subclasses
  - DIT (Depth of Inheritance Tree) = max_leng
    where, max_leng is the maximum length from the root node to the leaf node
Quality Measure Metrics

- To provide quality measurements
  - MHF (Method Hiding Factor) = \( \frac{\sum_{i=1}^{TC} \sum_{m=1}^{Md(C_i)} (1 - V(M_{mi}))}{\sum_{i=1}^{TC} Md(C_i)} \)

  where, \( V(M_{mi}) = \frac{\sum_{j=1}^{TC} is\_visible(M_{mi}, C_j)}{TC - 1} \)

  \( is\_visible(M_{mi}, C_j) = \begin{cases} 
  1 & \text{iff } j \neq i \text{ and } C_j \text{ may call } M_{mi} \\
  0 & \text{otherwise}
\end{cases} \)

  TC = Total number of class
  Md = Total number of methods defined
  V(M_{mi}) = Visibility of the total classes from which the method M_{mi} is visible

- MHF is a measure of the use of information hiding concept through methods
Quality Measure Metrics (Continued)

- AHF (Attribute Hiding Factor) = \[ \frac{\sum_{i=1}^{TC} \sum_{m=1}^{Ad} (1 - V(A_{mi}))}{\sum_{i=1}^{TC} Ad(C_i)} \]

where, \( V(A_{mi}) = \frac{\sum_{j=1}^{TC} is\_visible(A_{mi}, C_j)}{TC - 1} \)

\( is\_visible(A_{mi}, C_j) = \begin{cases} 1 & \text{iff } j \neq i \\ 0 & \text{otherwise} \end{cases} \)

TC = Total number of class
Ad = Total number of attributes defined
\( V(A_{mi}) = \) Visibility of the total classes from which the attribute \( A_{mi} \) is visible

- AHF is a measure of the use of information hiding concept through attributes
Quality Measure Metrics (Continued)

- **MIF (Method Inheritance Factor)** = \[ \frac{\sum_{i=1}^{TC} M_i(C_i)}{\sum_{i=1}^{TC} M_a(C_i)} \]
  where, \( M_a(C_i) = M_d(C_i) + M_i(C_i) \) is total number of available methods (locally defined plus inherited)
  - MIF is a measure of inheritance through methods

- **AIF (Attribute Inheritance Factor)** = \[ \frac{\sum_{i=1}^{TC} A_i(C_i)}{\sum_{i=1}^{TC} A_a(C_i)} \]
  where, \( A_a(C_i) = A_d(C_i) + A_i(C_i) \) is total number of available attributes (locally defined plus inherited)
  - AIF is a measure of inheritance through attributes
Coupling Metrics

- To provide dependency between objects/classes and locality of data
  - PLC (Package Level Coupling)
  - CLC (Class Level Coupling)
  - OLC (Object Level Coupling)

<Example of PLC>
Use Case Metrics

- To provide **dynamic complexity** through the analysis of **use case data**
  - NOA(Number of Actor) = $\sum_{i=1}^{n} noa_i$
  - NOUC(Number of Use Cases) = $\sum_{i=1}^{n} nouc_i$
  - NOUCA(Use Cases per Actor) = $\sum_{i=1}^{n} nouca_i$
Tool Support

<Structure of the OSMAT software>
Screen shots

- Sample results of running the OSMAT
A sample XML output

```xml
<?xml version="1.0" encoding="UTF-8" ?>
<!DOCTYPE Metrics (View Source for full doctype... )>
  - <Metrics>
    - <Primitive>
      <CLS TNC="0" TNP="9" TNP="0" TNC="12" />
      <REL TNR="0" TNUR="0" TNP="0" TNA="12" />
    </Primitive>
    - <Fault_Proneness DIT="0">
      <WMC ClassName="SSServer">0</WMC>
      <WMC ClassName="DataBase">0</WMC>
      <WMC ClassName="POSTMenu">0</WMC>
      <WMC ClassName="PayloadDefinition">0</WMC>
      <WMC ClassName="PayloadApplicationDefinition">0</WMC>
      <WMC ClassName="ADefineBasicPayloadInformation">1</WMC>
      <WMC ClassName="FDefineBasicPayloadInformation">8</WMC>
      <WMC ClassName="AuditData">0</WMC>
      <NOC ClassName="SSServer">0</NOC>
      <NOC ClassName="DataBase Server">0</NOC>
      <NOC ClassName="POSTMenu">0</NOC>
      <NOC ClassName="PayloadDefinition">0</NOC>
      <NOC ClassName="PayloadApplicationDefinition">0</NOC>
      <NOC ClassName="ADefineBasicPayloadInformation">0</NOC>
      <NOC ClassName="FDefineBasicPayloadInformation">0</NOC>
      <NOC ClassName="AuditData">0</NOC>
    </Fault_Proneness>
    <Quality_Measurement: MIF="0" MIF="0" MIF="0" MIF="0" />
    - <UseCase: Measurement NOCA="1" NOUCA="0">
      <NOUCA ActorName="User">0</NOUCA>
    </UseCase: Measurement>
  </Metrics>
```
Experimental Results

- Feasibility test
  - CDT (Command and Data Tool) software for Payload Operations Support Team tool suite
    - Command and telemetry specification for Space Shuttle payloads and experiments
    - Java-language client-server application
    - Shipped to Space Shuttle payload customers
  - 30 UML models
    - Used RUP, UML and variety of tools
    - Collected back-end metrics
    - Wanted front-end metrics
      - Can we determine complexity earlier?
      - Use measures for project decision-making
Experimental Results (Continued)

- CDT Application Example
### Experimental Results (Continued)

#### Results of running the OSMAT software

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic</td>
<td>TNC</td>
<td>8</td>
<td>17</td>
<td>21</td>
<td>16</td>
<td>21</td>
<td>8</td>
<td>13</td>
<td>12</td>
<td>13</td>
<td>30</td>
<td>21</td>
<td>37</td>
<td>51</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>TNIR</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>TNUR</td>
<td>0</td>
<td>20</td>
<td>28</td>
<td>13</td>
<td>15</td>
<td>14</td>
<td>32</td>
<td>11</td>
<td>9</td>
<td>53</td>
<td>24</td>
<td>63</td>
<td>15</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>TNRR</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>TNO</td>
<td>9</td>
<td>147</td>
<td>32</td>
<td>11</td>
<td>9</td>
<td>24</td>
<td>66</td>
<td>10</td>
<td>19</td>
<td>63</td>
<td>47</td>
<td>62</td>
<td>49</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>TNCA</td>
<td>73</td>
<td>178</td>
<td>200</td>
<td>49</td>
<td>94</td>
<td>91</td>
<td>262</td>
<td>83</td>
<td>40</td>
<td>306</td>
<td>217</td>
<td>204</td>
<td>110</td>
<td>141</td>
</tr>
<tr>
<td></td>
<td>TNA</td>
<td>6</td>
<td>7</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>11</td>
<td>18</td>
<td>27</td>
<td>6</td>
<td>103</td>
</tr>
<tr>
<td></td>
<td>TNR</td>
<td>12</td>
<td>14</td>
<td>0</td>
<td>2</td>
<td>10</td>
<td>0</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>22</td>
<td>36</td>
<td>54</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>210</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fault-</td>
<td>DIT</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Proneness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality</td>
<td>MIF</td>
<td>0.5000</td>
<td>0.5000</td>
<td>0.5000</td>
<td>0.5000</td>
<td>0.4375</td>
<td>0.5000</td>
<td>0.4527</td>
<td>0.5000</td>
<td>0.5000</td>
<td>0.5000</td>
<td>0.4960</td>
<td>0.5000</td>
<td>0.5000</td>
<td>0.5000</td>
</tr>
<tr>
<td></td>
<td>MIF</td>
<td>0.5000</td>
<td>0.5000</td>
<td>0.5000</td>
<td>0.5000</td>
<td>0.4375</td>
<td>0.5000</td>
<td>0.4527</td>
<td>0.5000</td>
<td>0.5000</td>
<td>0.5000</td>
<td>0.4960</td>
<td>0.5000</td>
<td>0.5000</td>
<td>0.5000</td>
</tr>
<tr>
<td></td>
<td>MIF</td>
<td>0.5000</td>
<td>0.5000</td>
<td>0.5000</td>
<td>0.5000</td>
<td>0.4375</td>
<td>0.5000</td>
<td>0.4527</td>
<td>0.5000</td>
<td>0.5000</td>
<td>0.5000</td>
<td>0.4960</td>
<td>0.5000</td>
<td>0.5000</td>
<td>0.5000</td>
</tr>
<tr>
<td></td>
<td>AF</td>
<td>0.4739</td>
<td>0.4739</td>
<td>0.4739</td>
<td>0.4739</td>
<td>0.4739</td>
<td>0.4739</td>
<td>0.4739</td>
<td>0.4739</td>
<td>0.4739</td>
<td>0.4739</td>
<td>0.4739</td>
<td>0.4739</td>
<td>0.4739</td>
<td>0.4739</td>
</tr>
<tr>
<td>Others</td>
<td>NOA</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>NOUC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

---

Copyright © 2003 TAMU
Utilization of the OSMAT

Cost Estimation Model

Generated by SAS with linear regression test

<table>
<thead>
<tr>
<th>Parameter Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>11</td>
</tr>
<tr>
<td>12</td>
</tr>
<tr>
<td>13</td>
</tr>
<tr>
<td>14</td>
</tr>
<tr>
<td>15</td>
</tr>
<tr>
<td>16</td>
</tr>
<tr>
<td>17</td>
</tr>
<tr>
<td>18</td>
</tr>
<tr>
<td>19</td>
</tr>
<tr>
<td>20</td>
</tr>
<tr>
<td>21</td>
</tr>
<tr>
<td>22</td>
</tr>
<tr>
<td>23</td>
</tr>
<tr>
<td>24</td>
</tr>
<tr>
<td>25</td>
</tr>
<tr>
<td>26</td>
</tr>
<tr>
<td>27</td>
</tr>
<tr>
<td>28</td>
</tr>
<tr>
<td>29</td>
</tr>
<tr>
<td>30</td>
</tr>
<tr>
<td>31</td>
</tr>
<tr>
<td>32</td>
</tr>
<tr>
<td>33</td>
</tr>
</tbody>
</table>

Model Equation

\[
\text{TOT} = 3091.31 - 148.363 \times \text{TNC} + 20.7095 \times \text{TNUR} - 546.326 \times \text{TNRR}
\]

- 5.5201 \times \text{TNO} + 17.0502 \times \text{TNCA} - 2031.25 \times \text{TNA} + 997.562 \times \text{TNR}

+ 982.287 \times \text{DIT} - 17822.3 \times \text{MIP} + 21811.3 \times \text{AHF} + 41021.4 \times \text{MIF}

- 3161.79 \times \text{AIP} + 2613.69 \times \text{NOA} + 21.7577 \times \text{NOUC} + 77.9870 \times \text{TNP}

+ 59.6933 \times \text{TNIR} - 0.6004 \times \text{P.18} - 9.5140 \times \text{P.19} - 0.0337 \times \text{P.20}

+ 13.0528 \times \text{P.21} - 2063.36 \times \text{P.22} + 1.6016 \times \text{P.25} - 0.4766 \times \text{P.26}

+ 521.317 \times \text{P.27} + 0.0072 \times \text{P.28} - 3.1399 \times \text{P.29} - 83662.2 \times \text{P.30}

+ 59665.8 \times \text{P.31} - 80324.5 \times \text{P.32}

Summary of Fit

<table>
<thead>
<tr>
<th>Mean of Response</th>
<th>R-Square</th>
<th>Adj R-Sq</th>
</tr>
</thead>
<tbody>
<tr>
<td>781.6807</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>Root MSE</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Utilization of the OSMAT (Continued)

- Correlation between metrics by T-test

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Pearson Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Negative</td>
</tr>
<tr>
<td>TNC</td>
<td>AIF (-0.506)</td>
</tr>
<tr>
<td>TNIR</td>
<td>TNA (-0.172)</td>
</tr>
<tr>
<td>TNUR</td>
<td>NOA (-0.221)</td>
</tr>
<tr>
<td>TNRR</td>
<td>TNCA (-0.116)</td>
</tr>
<tr>
<td>TNO</td>
<td>AIF (-0.474)</td>
</tr>
<tr>
<td>TNP</td>
<td>AIF (-0.583)</td>
</tr>
<tr>
<td>TNCA</td>
<td>MIF (-0.549)</td>
</tr>
<tr>
<td>TNA</td>
<td>MIF (-0.338)</td>
</tr>
<tr>
<td>TNR</td>
<td>MIF (-0.334)</td>
</tr>
<tr>
<td>DIT</td>
<td>TNO (0.339)</td>
</tr>
<tr>
<td>MHF</td>
<td>TOT (-0.186)</td>
</tr>
<tr>
<td>AHF</td>
<td>AIF (-0.602)</td>
</tr>
<tr>
<td>MIF</td>
<td>TNCA (-0.549)</td>
</tr>
<tr>
<td>AIF</td>
<td>TNP (-0.583)</td>
</tr>
<tr>
<td>NOA</td>
<td>TNUR (-0.221)</td>
</tr>
<tr>
<td>NOUC</td>
<td>DIT (-0.165)</td>
</tr>
<tr>
<td>TOT</td>
<td>TNO (-0.245)</td>
</tr>
</tbody>
</table>
Summary

- What Metrics?
  - Object-Oriented Metrics to understand software complexity
  - Propose new metrics based on Use Case diagrams

- How to collect?
  - Automatic Measurement Approach: Ontology-based Software Metrics Analysis Tool

- How to utilize?
  - Effort estimation model
  - Statistical analyses: Regression, Cluster
Contact Information

- Peter In
  
  Assistant Professor
  Dept of Computer Science
  Texas A&M University
  College Station, TX 77843-3112

  Email: hohin@cs.tamu.edu
  Web: http://faculty.cs.tamu.edu/hohin
  Voice: +1–979-458-1547