COCOTS
Software Integration Cost Model: Insights & Status

Dr. Barry Boehm (USC)
Mr. Chris Abts (USC)
Dr. Betsy Bailey (Software Metrics)

USC Center for Software Engineering
Annual Research Review, Los Angeles, CA
February 8, 1999
COTS Integration Cost Sources:

1) Assessment

Initial Filtering Effort

Total Effort = \( \left( \text{# COTS Candidates} \right) \left( \frac{\text{Average Filtering Effort}}{\text{Candidate}} \right) \)

Final Selection Effort

Total Effort = \( \sum_{\text{Assessment Attributes}} \left( \text{# COTS Candidates} \right) \left( \frac{\text{Average Assessment Effort for Attribute in Given Domain}}{\text{Candidate}} \right) \)

- List of attributes refined in collaboration with Dr. Elizabeth Bailey
- Effort/candidate is project-dependent, within domain guidelines
COTS Integration Cost Sources:

2) Tailoring

Total Effort = \( \sum_{i} \left( \frac{\text{# COTS Candidates Tailored at Complexity Level}}{\text{Average Effort at Tailoring Complexity Level in Domain}} \right) \)

- Five tailoring effort complexity levels:
  - Very Low, Low, Nominal, High, Very High
  - Differentiated based on number tailored parameters, difficulty of needed scripts, API iterations, etc.

COTS Integration Cost Sources:

3) Glue Code Development and Test

Total Effort = \( A \cdot \left[ \text{size} \cdot (1 + \text{breakage}) \right] \) (effort multipliers)

- \( A \) - a linear scaling constant
- \( \text{Size} \) - of the glue code in SLOC or FP
- \( \text{Breakage} \) - of the glue code due to change in requirements and/or COTS volatility
- \( \text{Effort Multipliers} \) - 13 parameters, each with settings ranging VL to VH
- \( B \) - an architectural scale factor with settings VL to VH
**COTS Integration Cost Sources:**

4) Increased Application Effort Due to COTS Volatility

**Approximate Model:**

Total Effort = (Application Effort) \times \left[ \frac{\text{BRAK COTS}}{100} \right] \times \text{(EAF)}

**Detailed Model with COCOMO II Parameters:**

Total Effort = (Application Effort) \times \left( \frac{1 - \text{BRAK COTS}}{1 + \text{BRAK}} \right)^{1.01 + \Sigma} \times \text{(EAF)}

- **BRAK COTS:** % application code breakage due to COTS volatility
- **BRAK:** % application code breakage otherwise
- **\Sigma:** COCOMO II scale factor
- **EAF:** Effort Adjustment Factor (product of effort multipliers)

---

**Total COTS Integration Cost Estimate**

Total Integration Effort (in Person-Months) =

Assessment Effort + Tailoring Effort + Glue Code Effort + Volatility Effort

where

Assessment Effort = Filtering Effort + Final Selection Effort

Total integration Cost =

\[(\text{Total Integration Effort}) \times \text{(SS/Person-Month)}\]
Data Collection Status

• 6 Student Digital Library Projects
  – 8 more by end Spring ‘99 semester

• 12 Industrial Projects
  – FAA & aerospace contractors
  – 8+ additional projects anticipated by mid ‘99
  – will allow calibration of Early Design version

• Other Sources Being Explored
  – NASA, DoD, Commercial
  – USC-CSE Affiliates, GSAW & ICSE conferences

Experiences with Student Data

Highlights

• Raw Data
• COTS Assessment Effort Distribution Profile
• Glue Code Submodel Calibration Result
• Insights from Student Projects
Experiences with Student Data

Raw Project Data

<table>
<thead>
<tr>
<th>ACTIVITY</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Group 4</th>
<th>Group 5</th>
<th>Group 6</th>
<th>Total Person-Hrs</th>
<th>% Total Person-Hrs by Activity</th>
<th>% Total Person-Hrs by Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Activity:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Determine Requirements:</td>
<td>15.00</td>
<td>49.50</td>
<td>40.50</td>
<td>26.50</td>
<td>5.50</td>
<td>38.50</td>
<td>232.50</td>
<td>6.99</td>
<td>4.99</td>
</tr>
<tr>
<td>Prepare update plans:</td>
<td>157.00</td>
<td>142.50</td>
<td>269.50</td>
<td>36.50</td>
<td>123.50</td>
<td>154.75</td>
<td>757.75</td>
<td>10.06</td>
<td>10.06</td>
</tr>
<tr>
<td>Design project:</td>
<td>90.00</td>
<td>3.00</td>
<td>103.50</td>
<td>53.50</td>
<td>13.50</td>
<td>18.00</td>
<td>378.00</td>
<td>8.43</td>
<td>8.43</td>
</tr>
<tr>
<td>Code project:</td>
<td>151.00</td>
<td>20.50</td>
<td>190.00</td>
<td>146.50</td>
<td>67.50</td>
<td>115.50</td>
<td>722.00</td>
<td>15.23</td>
<td>15.23</td>
</tr>
<tr>
<td>Participate in formal design/code reviews:</td>
<td>14.00</td>
<td>9.00</td>
<td>21.00</td>
<td>21.00</td>
<td>22.50</td>
<td>24.00</td>
<td>110.00</td>
<td>2.43</td>
<td>2.43</td>
</tr>
<tr>
<td>Integrate and test:</td>
<td>73.00</td>
<td>84.50</td>
<td>85.50</td>
<td>6.50</td>
<td>13.50</td>
<td>29.50</td>
<td>299.00</td>
<td>6.71</td>
<td>6.71</td>
</tr>
<tr>
<td>Fix defects found in testing:</td>
<td>83.00</td>
<td>27.50</td>
<td>61.00</td>
<td>2.00</td>
<td>15.00</td>
<td>71.00</td>
<td>296.50</td>
<td>5.14</td>
<td>5.14</td>
</tr>
<tr>
<td>COTS Related Activity:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Understand and qualify COTS:</td>
<td>2.00</td>
<td>6.00</td>
<td>38.50</td>
<td>10.00</td>
<td>61.00</td>
<td>19.50</td>
<td>197.00</td>
<td>4.22</td>
<td>4.22</td>
</tr>
<tr>
<td>Design COTS glue codes:</td>
<td>0.00</td>
<td>5.00</td>
<td>7.50</td>
<td>0.00</td>
<td>5.00</td>
<td>9.00</td>
<td>16.00</td>
<td>0.38</td>
<td>0.38</td>
</tr>
<tr>
<td>Code COTS glue codes:</td>
<td>0.00</td>
<td>0.00</td>
<td>4.00</td>
<td>0.00</td>
<td>16.00</td>
<td>30.50</td>
<td>51.50</td>
<td>1.15</td>
<td>1.15</td>
</tr>
<tr>
<td>Fix defects found in COTS testing:</td>
<td>5.00</td>
<td>0.00</td>
<td>2.50</td>
<td>1.00</td>
<td>1.50</td>
<td>4.00</td>
<td>14.00</td>
<td>0.31</td>
<td>0.31</td>
</tr>
<tr>
<td>Administrative Activity:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Management:</td>
<td>8.50</td>
<td>34.00</td>
<td>43.00</td>
<td>15.00</td>
<td>10.00</td>
<td>21.50</td>
<td>124.50</td>
<td>2.78</td>
<td>2.78</td>
</tr>
<tr>
<td>Documentation:</td>
<td>52.30</td>
<td>445.00</td>
<td>36.50</td>
<td>59.00</td>
<td>66.50</td>
<td>126.00</td>
<td>762.50</td>
<td>17.61</td>
<td>17.61</td>
</tr>
<tr>
<td>Other:</td>
<td>114.00</td>
<td>239.00</td>
<td>31.50</td>
<td>8.00</td>
<td>100.00</td>
<td>82.50</td>
<td>575.00</td>
<td>12.90</td>
<td>12.90</td>
</tr>
<tr>
<td>TOTAL WEEKLY Person-Hours:</td>
<td>705.00</td>
<td>1073.50</td>
<td>972.50</td>
<td>415.50</td>
<td>477.50</td>
<td>699.25</td>
<td>4450.35</td>
<td>95.98</td>
<td>95.98</td>
</tr>
</tbody>
</table>

Table VIII-1: Effort hours by activity for graduate software engineering class projects incorporating COTS products.

UNIVERSITY OF SOUTHERN CALIFORNIA

Experiences with Student Data

COTS Assessment Effort Distribution

Groups 3 & 5 (search engines)

<table>
<thead>
<tr>
<th>Gross Attributes</th>
<th>Activities I</th>
<th>Activities II</th>
<th>Activities III</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Functionality</td>
<td>20%</td>
<td>10%</td>
<td>30%</td>
</tr>
<tr>
<td>2. Performance</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
</tr>
<tr>
<td>3. Dependability</td>
<td>10%</td>
<td>10%</td>
<td>20%</td>
</tr>
<tr>
<td>4. Usability</td>
<td>10%</td>
<td>10%</td>
<td>20%</td>
</tr>
<tr>
<td>5. Adaptability</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
</tr>
<tr>
<td>6. Operability</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
</tr>
<tr>
<td>7. Cost</td>
<td>50%</td>
<td>20%</td>
<td>30%</td>
</tr>
</tbody>
</table>

Distribution of assessment effort by activity and attribute:

I: nominal exercise - use COTS as intended by vendor
II: off-nominal exercise - adapt COTS to new use
III: reading and research
Experiences with Student Data
Glue Code Submodel Calibration

<table>
<thead>
<tr>
<th>Project</th>
<th>A (SLOC)</th>
<th>B</th>
<th>xEAFs</th>
<th>Estimate (P-hr)</th>
<th>Actual (P-hr)</th>
<th>Relative Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.009</td>
<td>500</td>
<td>1.04</td>
<td>1.82</td>
<td>10.50</td>
<td>-9%</td>
</tr>
<tr>
<td>5</td>
<td>0.009</td>
<td>400</td>
<td>1.12</td>
<td>2.25</td>
<td>16.62</td>
<td>-3%</td>
</tr>
<tr>
<td>6</td>
<td>0.009</td>
<td>218</td>
<td>1.16</td>
<td>10.42</td>
<td>48.38</td>
<td>22%</td>
</tr>
</tbody>
</table>

A = .009 => 111 SLOC/P-hr

Insights from Student Data

- Like Student COCOMO, there is utility in developing a Student COCOMO
  - scaling issues (SLOC/P-hr vs. KSLOC/P-mth)
  - student raw productivity higher (don’t have same security, overhead, coordination, documentation, version control, etc., concerns as industrial developers)

- Helping us to sort out what is COTS specific vs. COCOMO specific effort

- Illustrates again efficacy of a site-specific calibration
Experience with Data Interviews

• lessons learned
• modeling suggestion:

Apply model by COTS class rather than at component or project level
(good compromise for data collection?)

Suggested COTS Classes

• database
• network management
• GUI builders
• operating systems
• report generators
• device drivers
• compilers
• decision support systems
• other???
**Immediate COCOTS Follow-ons**

- Modeling of schedule estimation and activity distribution
- Integration with COCOMO II estimation model
- More extensive tool implementation

**Conclusions**

- COCOTS is still evolving/defining its framework for estimating software COTS integration and usage costs
  - Data collection interviews adding immeasurably to our insight into COTS integration, allowing capture of unique "lessons learned" that are helping to refine the model
  - Project database growing, soon will reach critical mass needed for first publishable calibration, at least of Early Design version
  - Important schedule/activity distribution features, initial formal reconciliation with COCOMO II model anticipated by end '99
- COCOTS can be extended to cover other COTS related costs
  - Biggest challenge will be complex, dynamic COTS price structures
Backup Slides

Outline

- Model Development History and Support
- Problem Context
- COTS Software Integration Cost Sources
- Early Design/Post-architecture Model Versions
- Longer-term COCOTS Follow-ons
Model Development History and Support

- USAF/ESC Effort
  - March 1996 through June 1997
    - Initial Glue Code Model Definition, Experimental Calibration

- FAA Effort
  - Phase 1 (July to October, 1997)
    - Glue Code Model Redefinition, Experimental Calibration
  - Phase 2 (October 1997 to July 1998)
    - Glue Code Model Refinement
    - Assessment, Tailoring, and Volatility Models Defined
  - Phase 3 (July 1998 to December 1998)
    - Further Data Collection & Model Refinement, Calibration
    - Goal: calibrated model available by end 1998

- ONR Effort
  - January 1998 through 1999
    - Further Refinement of Models, including activity analysis & effort distribution
    - Data Collection & Calibration
    - Determination of How Best to Associate COCOTS with COCOMO II

Problem Context: Modeling

COTS ■ and Custom C Applications Components

New COCOTS Modeling Problem

COTS Infrastructure
COCOMO II: PVOL, PEXP

COTS Tools
LTEX, TOOL

Cost Modeling Currently Addressed
COTS Integration Cost Sources:

1) Assessment - Assessment Attributes

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Codes</th>
<th>Functionality</th>
<th>Portability</th>
<th>Price</th>
<th>Interoperability</th>
<th>Recurring Costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Documentation Quality</td>
<td>DQ</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td></td>
</tr>
<tr>
<td>Correctness</td>
<td>C</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td></td>
</tr>
<tr>
<td>Availability</td>
<td>A</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>Fail safe</td>
<td>FS</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>Reliability</td>
<td>R</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>Redundancy</td>
<td>R</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>Security</td>
<td>S</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>Security related</td>
<td>SR</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>Product Performance</td>
<td>P</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>Information/Data capacity</td>
<td>IDC</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>Memory performance</td>
<td>MP</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>Response time</td>
<td>RT</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>Throughput</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>Testability</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>Locality</td>
<td>L</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>Vendor Support</td>
<td>V</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>Willingness to escrow source code</td>
<td>WESC</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>Willingness to make modifications</td>
<td>WMOD</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
</tbody>
</table>

2) Tailoring - Dimensions of Tailoring Difficulty

- Difficulty Table
- Vendor Code
- Initial Purchase
- Recurring Costs
- Vendor Support
- User Training
- Willingness to escrow source code
- Willingness to make modifications
COTS Integration Cost Sources:
3) Glue Code Development and Test - Glue Code Cost Drivers

Personnel Drivers
1) ACIEP - COTS Integrator Experience with Product
2) ACIPC - COTS Integrator Personnel Capability
3) AXCICP - Integrator Experience with COTS Integration Processes
4) APCR - Integrator Personnel Continuity

COTS Component Drivers
5) ACPCT - COTS Product Maturity
6) ACPX - COTS Supplier Product Extension Willingness
7) ACPX - COTS Product Interface Complexity
8) ACPXS - COTS Supplier Product Support
9) ACPCTD - COTS Supplier Provided Training and Documentation

Application/Systems Drivers
10) ACRE - Constraints on Application System/Subsystem Reliability
11) AACPX - Application Interface Complexity
12) ACPER - Constraints on COTS Technical Performance
13) ASPRT - Application System Portability

Nonlinear Scale Factor
1) AAREN - Application Architectural Engineering

Recent Development: two models, differing fidelity
(Parallels COCOMO II modeling)

Early Design COCOTS model
- roll up of parameters in Assessment, Glue code submodels into fewer, more aggregated factors; inclusion of only the approximate Volatility model.
- less fidelity but requires fewer data points to calibrate.
- intended for more “what if” kind of estimating, earlier in the development process.

Post-architecture COCOTS model
- the full model as presented in preceding charts
Longer-term COCOTS Follow-ons

- Continued data collection and conditioning
- Continued recalibration and iteration of the model within current structure
- Experimental usage and refinement, including exploration of other cost drivers and model forms
- Modeling other COTS related costs
  - Licenses, training, maintenance, hardware