Early COCOTS – A Second Iteration

Betsy Clark
Brad Clark
Software Metrics, Inc.
October 26, 2004

Presentation Outline

- Background
  - COCOTS
  - Early COCOTS (First iteration)
- Reality Check with Model Users
- Reformulation
  - Early COCOTS (Second iteration)
- Next Steps
What is COCOTS?

- Constructive Cost Model for systems containing COTS components
- Part of the COCOMO suite
- Estimates effort for
  - Assessment
  - Tailoring
  - Glue Code development
- Addresses initial development as well as maintenance
- Calibrated with data collected from 21 development projects (plus maintenance data for some of these)

Background

- From trial applications of COCOTS, two problems became clear
  1. Many of the cost drivers are important but are not typically known early
     - Example: Number of screens and reports
     - People have a difficult time estimating these until they are well into prototyping or operational evaluations.
  2. Major sources of cost were missing altogether (e.g., BPR)
- A model is needed that can be used very early on in the life cycle to address basic investment decisions
- The challenge is to create an early estimating model based on what people actually know (or can reason about)
What is Early COCOTS?

- 35,000 foot view to be used early in the lifecycle for:
  - rough order-of-magnitude (ROO) estimates
- Simplified model
  - Information known early in the lifecycle is limited

Objectives for Early COCOTS

- Should handle COTS, NDI, and new code
  - Entire range of possibilities
- Cost drivers can be estimated or are known early on
- Model should give a range of outputs, not a single point estimate
- Costs will be estimated at the system level, not at the level of individual components
- Model should address total cost of ownership
  - Actual cost of the COTS products
  - Effort to build / integrate the COTS-Based system
  - BPR, training, consulting
  - Maintenance costs
(More) Background

- Brad and Betsy put together a strawman set of cost drivers
- Barry Boehm reworked them and created rating scales
  - Workshop led by Brad at COCOMO forum in 2002

Early COCOTS – First Iteration

Sizing
- Number of independent COTS products
- Amount of newly developed software (equivalent SLOC)
- % functions provided by COTS
- Degree of uncertainly about product choices
Early COCOTS – First Iteration

Effort Multipliers
- Complexity of Integration
  - Required tailoring
  - Data conversion
  - Business Processing Reengineering, training,
  - Integrator difficulties with COTS products and integration
- Degree of mismatch between COTS capabilities and user needs; maturity, requirements flexibility
- Requirements Volatility
- COTS Volatility

Reality Check – Involving Model Users
- We thought we had a good set of cost drivers but we’re not the ones who will be using the model
  - We didn’t want to develop this model in a vacuum
- Within the FAA, primary cost estimation responsibility and expertise resides within an Investment Analysis group
- By getting them involved in reviewing the model, we hoped for buy-in (and a better model!)
FAA Workshops

- Held two workshops: March, April 2004
- Focus
  - Do we have a reasonable set of early lifecycle cost drivers?
  - Are we missing any?
  - Do the scales make sense?
    - i.e., can we improve on the descriptions?
- Reviewed model components
  - Size inputs
  - Effort multipliers
- We expected some model "tweaking"
  - We were in for a surprise!
- One of the participants came to the second workshop with a list of what information is known during Investment Analysis
  - List was further refined
  - This functioned as a "wish list" of areas a cost model should consider

Investment Analysis

- High-level:
  - Requirements defined
  - Model: Early COTS
- Developer/Integrator known
- COTS products known
- Model: COTS
What Model Users Know Early On

- System Sizing Parameters
  - Number of users
  - Number of sites
  - Amount of data (legacy and new)
  - Number and age of databases to be converted
  - Amount of legacy code to be reused
  - Totally new systems are rare
  - Number of interfacing systems

- Requirements
  - Functional (high level)
  - Performance
  - Security

- Architecture (Solution alternatives)
- Implementation Strategy

What Model Users Don't Know

- Who the contractor will be
  - They know who the likely bidders will be
- Specific COTS products
- Precise number of COTS products
FAA Workshop Follow-up

- Discussed with USC: June 2004
  - Many of these items in the list are system engineering concerns
  - Scoping issue
    - Too much to expect Early COCOTS to address system characteristics as well
  - These areas are addressed by COSYSMO
  - Led to integrated solution with COSYSMO, Early CCCOTS, and Early Design COCOMO II
    - With Early COCOTS becoming really simple

Status: Reformulated Approach

- Changed approach to "multi-model" for estimating COTS-based system costs
- For a software system of interest:
  - System engineering costs for software-intensive systems are estimated by the emerging COSYSMO model
  - Custom development portion of the software system estimated by COCOMO II Early-Design
  - COTS based portion of the software system estimated by Early-COCOTS

\[
\text{Estimated System Cost} = \text{COSYSMO} + \text{Early Design COCOMO II} + \text{Early COCOTS}
\]
### COSYSMO Effort Drivers

- Requirements Understanding
- Architecture Understanding
- Level-of-Service Requirements
- Migration Complexity
- Technology Maturity
- Documentation Match to Life Cycle Needs
- Diversity of Installations/Platforms
- Recursive Levels in the Design
- Stakeholder Team Cohesion
- Personnel/Team Capability
- Personnel Experience/Continuity
- Process Capability
- Multisite Coordination

---

### Size Drivers

**COSYSMO**

- # system requirements
- # interfaces
- # of critical algorithms
- # of operational scenarios

**Early COCOTS – Second Iteration**

- # critical COTS products
- Degree of uncertainty about COTS product choices

**Early COCOMO II**

- Amount of custom code (ESLOC)
Early COCOTS – Second Iteration

- KISS – to the max!
- Addresses Assessment, Tailoring, Glue Code effort
- For each of these activities, our goal (and challenge) is to identify simple effort drivers that are known early on
- Let's look at two possibilities for Assessment:
  - (1) Number of products being assessed
  - (2) Degree of Uncertainty about Product Choice

Assessment Effort vs. Number of COTS Products

![Graph showing Assessment Effort vs. Number of COTS Products]
Uncertainty about Product Choice

- Rating scale in Early COCOTS – First Iteration

<table>
<thead>
<tr>
<th>Driver</th>
<th>Small</th>
<th>Medium</th>
<th>Large</th>
<th>Very Large</th>
<th>Extra Large</th>
</tr>
</thead>
<tbody>
<tr>
<td>Degree of Uncertainty about product choice</td>
<td>Clear choices</td>
<td>Manual ratings of key criteria</td>
<td>Simple exercise of 1 to 3 candidates</td>
<td>Thorough benchmarking, prototyping to access key criteria</td>
<td>Major evaluations of complex interoperability</td>
</tr>
</tbody>
</table>

Assessment Effort vs Degree of Product Uncertainty

![Graph showing assessment effort vs uncertainty]

October 26, 2004 21

October 26, 2004 22
Assessment Input

The effort required for Assessment is by the Degree of Uncertainty.

Select the Degree of Uncertainty

<table>
<thead>
<tr>
<th></th>
<th>Small</th>
<th>Medium</th>
<th>Large</th>
<th>Very Large</th>
<th>Extra Large</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower 80% CL</td>
<td>0.31</td>
<td>0.78</td>
<td>2.27</td>
<td>5.16</td>
<td>11.65</td>
</tr>
<tr>
<td>Mean</td>
<td>0.37</td>
<td>1.00</td>
<td>2.37</td>
<td>7.44</td>
<td>20.14</td>
</tr>
<tr>
<td>Upper 80% CL</td>
<td>0.44</td>
<td>1.27</td>
<td>3.28</td>
<td>10.73</td>
<td>34.38</td>
</tr>
</tbody>
</table>

Estimated Assessment Effort = 8.27 PM * Uncertainty Rating Range

Estimating Assessment Example:
The Degree of Uncertainty was judged as Large (using rating descriptions or analogy data to be provided)
Low Estimate = 8.27 PM * 2.27 = 19 PM
Mean Estimate = 8.27 PM * 2.73 = 23 PM
High Estimate = 8.27 PM * 3.28 = 27 PM

Degree of Uncertainty -1

- Low
  - Select from list of pre-certified products
  - Choice is dictated
  - Already using a product which will be used in this project
- Medium
  - There are multiple products but not detailed assessment. Simple exercising of the products, paper and pencil evaluation
- Large
  - One or two products got very detailed assessment and the other products choices were certain, e.g. once the operating system was chosen the other products were selected as well
Degree of Uncertainty: -2

- **Very Large**
  - Fair number of COTS products with very high level of service requirements combined large amount of custom code. There is a lot of uncertainty and risk around those products. A lot of effort is spent on making sure those products work.
  - Verify service level agreements such as performance, reliability, availability, fault tolerance, security, interoperability, etc.
  - Quadruple redundancy
  - Seven 9’s of reliability (99.99999)
  - Through benchmarking, prototyping to assess key criteria

- **Extra Large**
  - Many different groups of users: end-to-end detailed scenarios (entire work flows and data flows) required to assess suitability
  - Example: Government Financial package suite used for multiple government agencies

---

Current Analysis

- Currently working on Tailoring and Glue Code
- Again, our goal and challenge is to find a simple effort driver that jumps out from the data and can be known early on
Mapped COCOTS Data to COSYSMO Cost Drivers

- **One hypothesis**: COTS-unique effort (assessment, tailoring, glue code) is affected by the COSYSMO drivers
- But, at least so far, nothing jumps out at us as a major driver of COTS-related effort

Areas to be Addressed

- Model should address total cost of ownership
  - Actual cost of the COTS products
  - Effort to build / integrate the COTS-Based system
  - BPR, training, consulting
  - Maintenance costs
Training Cost Estimate Example

<table>
<thead>
<tr>
<th>Sample Activity</th>
<th>Preparation for Training</th>
<th>Classroom Training</th>
<th>Periodic Training on new procedures</th>
<th>Software Development</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost Estimating Relationship (CER)</td>
<td>Activity-based</td>
<td>Unit costing</td>
<td>Analogy-based</td>
<td>Parametric</td>
</tr>
<tr>
<td>Rule</td>
<td>10-20 hours for each Class Hour</td>
<td>N trainers total M trainees</td>
<td>It cost us $XXX last year...</td>
<td>Early Design COCOMO II</td>
</tr>
</tbody>
</table>

October 26, 2004

Analogy Data

- Projects in COCOTS database will be described in terms of
  - application area
  - Early COCOTS drivers
  - COSYSMO drivers
- Investment Analysis team can use these analogous systems to help arrive at driver ratings

October 26, 2004
Results of this work

- Model Manual Definition
  - With guidelines for using
- Model Definition
- Analogy Data
- Data collection framework to extend model
  - e.g., for Business Process Reengineering
- Software tool (USC)

For More Information

Brad Clark
(703) 754-0115
Brad@software-metrics.com

Ye Yang
(213) 740-6470
yey@usc.edu

Betsy Clark
(703) 754-0115
Betsy@software-metrics.com

http://sunset.usc.edu

October 26, 2004