UNIVERSITY OF
SOUTHERN CALIFORNIA

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

UNIVERSITY OF
SOUTHERN CALIFORNIA

USC Center for Software Engineering

Points of Contact at USC-CSE in Los Angeles

Mr. Christopher Abts (primary graduate researcher).....(213) 740-6470
Ms. Ladonna Pierce (CSE Office Administrator).....(213) 740-5703
Dr. Barry W. Boehm (CSE Director).....(213) 740-8163
USC Center for Software Engineering FAX line.....(213) 740-4927
COCOTS E-Mail.....*cots-info@sunset.usc.edu*
World Wide Web page.....*http://sunset.usc.edu/COCOTS/cocots.html*

Additional Contact at Software Metrics, Inc. in Virginia (near D.C.)

Dr. Elizabeth (Betsy) Bailey.....(703) 754-0115
FAX line.....(703) 754-0115
E-Mail.....bkbailey@erols.com

UNIVERSITY OF SOUTHERN CALIFORNIA

Outline

- COTS Software Integration Cost Sources
- Data Collection Status
- Experiences with Student Data
- Experience with Data Interviews
- Immediate COCOTS Follow-ons
- Conclusions

UNIVERSITY OF SOUTHERN CALIFORNIA

COTS Integration Cost Sources:
1) Assessment

Initial Filtering Effort

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

Final Selection Effort

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

- *List of attributes refined in collaboration with Dr. Elizabeth Bailey*
- *Effort/candidate is project-dependent, within domain guidelines*

UNIVERSITY OF SOUTHERN CALIFORNIA

COTS Integration Cost Sources: 2) Tailoring

$$\text{Total Effort} = \sum_{\substack{\text{Tailoring} \\ \text{Complexity} \\ \text{Levels}}} \left(\begin{matrix}\text{\# COTS Candidates} \\ \text{Tailored at} \\ \text{Complexity Level}\end{matrix}\right)_i \left(\begin{matrix}\text{Average Effort at Tailoring} \\ \text{Complexity Level in Domain}\end{matrix}\right)_i$$

–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.

UNIVERSITY OF SOUTHERN CALIFORNIA

COTS Integration Cost Sources: 3) Glue Code Development and Test

$$\text{Total Effort} = A \cdot [(\text{size})(1+\text{breakage})]^{B} \cdot \prod (\text{effort multipliers})$$

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

UNIVERSITY OF SOUTHERN CALIFORNIA

COTS Integration Cost Sources: 4) Increased Application Effort Due to COTS Volatility

Approximate Model:

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

Detailed Model with COCOMO II Parameters:

$$\text{Total Effort} = (\text{Application Effort}) \cdot \left[\left(1+\frac{\text{BRAK COTS}}{1+\text{BRAK}}\right)^{1.01+\Sigma} - 1\right] \cdot (\text{EAF})_{\text{COTS}}$$

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

UNIVERSITY OF SOUTHERN CALIFORNIA

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 =

(Total Integration Effort) • ($$/Person-Month)

UNIVERSITY OF
SOUTHERN CALIFORNIA

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

UNIVERSITY OF
SOUTHERN CALIFORNIA

Experiences with Student Data
Highlights

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

UNIVERSITY OF SOUTHERN CALIFORNIA

Experiences with Student Data

Raw Project Data

ACTIVITY	Group 1 Edgar Cpr	Group 2 Med. Mscrp	Group 3 Tech Rpts	Group 4 LAPIS	Group 5 CNTV Archv	Group 6 Hancock PH	Total Pers-hrs by Activity	% Total Pers-hrs by Activity
General Activity								
Determine Requirements:	16.00	49.50	86.50	26.50	5.50	38.50	222.50	4.99
Prepare, update plans :	107.00	142.00	209.50	39.00	83.50	134.75	715.75	16.06
Design product :	99.00	3.00	103.50	63.50	13.00	96.00	378.00	8.48
Code product :	161.00	20.50	190.00	168.00	67.50	115.00	722.00	16.20
Participate in formal design/code reviews:	14.00	8.00	21.00	21.00	22.50	24.00	110.50	2.48
Integrate and test :	70.00	94.50	85.50	6.50	13.00	29.50	299.00	6.71
Fix defects found in testing:	60.00	27.50	61.00	2.00	15.00	71.00	236.50	5.31
COTS Related Activity								
Understand and qualify COTS:	2.00	6.00	98.50	10.00	61.00	19.50	197.00	4.42
Design COTS glue code :	0.00	0.00	7.50	0.00	0.30	9.00	16.80	0.38
Code COTS glue code :	0.00	0.00	4.00	0.00	16.80	30.50	51.30	1.15
Fix defects found in COTS testing:	5.00	0.00	2.50	1.00	1.50	4.00	14.00	0.31
Administrative Activity								
Management:	8.50	34.00	33.50	13.50	10.00	25.00	124.50	2.79
Documentation :	52.50	449.50	38.00	59.50	68.00	126.00	793.50	17.81
Other:	114.00	239.00	31.50	8.00	100.00	82.50	575.00	12.90
TOTAL WEEKLY Person-Hours	709.00	1073.50	972.50	418.50	477.60	805.25	4456.35	99.99

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

Key: Group 1 - EDGAR Corporate Data
Group 2 - Medieval Manuscripts
Group 3 - Technical Reports
Group 4 - Latin American Pamphlets
Group 5 - CNTV Moving Image Archive
Group 6 - Hancock Photo Archive

UNIVERSITY OF SOUTHERN CALIFORNIA

Experiences with Student Data

COTS Assessment Effort Distribution

Groups 3 & 5 (search engines)

Gross Attributes	Activities I	II	III	
1. Functionality	20%		10%	30%
2. Performance	10%			10%
3. Dependability	10%	10%		20%
4. Usability	10%		10%	20%
5. Adaptability		10%		10%
6. Operability			10%	10%
7. Cost				
	50%	20%	30%	100%

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

UNIVERSITY OF SOUTHERN CALIFORNIA

Experiences with Student Data

Glue Code Submodel Calibration

Project	A	Size (SLOC)	B	xEAFs	Estimate (P-hr)	Actual (P-hr)	Relative Error
3	0.009	500	1.04	1.82	10.50	11.50	-9%
5	0.009	400	1.12	2.25	16.62	17.10	-3%
6	0.009	218	1.16	10.42	48.38	39.50	22%

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

UNIVERSITY OF SOUTHERN CALIFORNIA

Insights from Student Data

- Like *Student COCOMO*, there is utility in developing a *Student COCOTS*
  - 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

UNIVERSITY OF
SOUTHERN CALIFORNIA

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?)

UNIVERSITY OF
SOUTHERN CALIFORNIA

Suggested COTS Classes

- database
- network management
- GUI builders
- operating systems
- report generators
- device drivers
- compilers
- decision support systems
- other???

UNIVERSITY OF
SOUTHERN CALIFORNIA

Immediate COCOTS Follow-ons

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

UNIVERSITY OF
SOUTHERN CALIFORNIA

Conclusions

- COCOTS is still evolving/refining 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

UNIVERSITY OF
SOUTHERN CALIFORNIA

Backup Slides

UNIVERSITY OF
SOUTHERN CALIFORNIA

Outline

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

UNIVERSITY OF SOUTHERN CALIFORNIA

Model Development History and Support

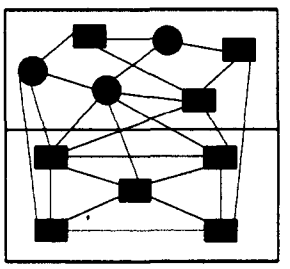
- USAF/ESC Effort
  - March 1996 through June 1997
    - Initial Glue Code Model Definition, Experimental Calibration
- FAA Effort
  - Phase 1 (July to October, 1997)
    - Glue Code Model Redefinition, Experimental Calibration
  - Phase 2 (October 1997 to July 1998)
    - Glue Code Model Refinement
    - 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

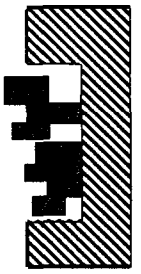
UNIVERSITY OF SOUTHERN CALIFORNIA

Problem Context: Modeling

COTS ■ and Custom● Applications Components

New **COCOTS** Modeling Problem





COTS Infrastructure
COCOMO II: PVOL, PEXP

COTS Tools
LTEX, TOOL

Cost Modeling Currently Addressed

UNIVERSITY OF SOUTHERN CALIFORNIA

COTS Integration Cost Sources:
1) Assessment - Assessment Attributes

Correctness
- Accuracy
- Correctness

Availability/Robustness
- Availability
- Fail safe
- Fail soft
- Fault tolerance
- Input error tolerance
- Redundancy
- Reliability
- Robustness
- Safety

Security
- Security (Access related)
- Security (sabotage related)

Product Performance
- Execution performance
- Information/data capacity
- Precision
- Memory performance
- Response time
- Throughput

Understandability
- Documentation quality
- Simplicity
- Testability

Ease of use
- Usability/Human Factors

Version Compatibility
- Downward compatibility
- Upward compatibility

Inter-component Compatibility
- Compatibility with other components
- Interoperability

Flexibility
- Extendability
- Flexibility

Installation/Upgrade Ease
- Installation Ease
- Upgrade/Refresh ease

Portability
- Portability

Functionality
- Functionality

Price
- Initial purchase/lease
- Recurring costs

Maturity
- Product Maturity
- Vendor Maturity

Vendor Support
- Response time for critical problems
- Support
- Warranty

User Training
- User training

Vendor Concessions
- Willingness to escrow source code
- Willingness to make modifications

UNIVERSITY OF SOUTHERN CALIFORNIA

COTS Integration Cost Sources:
2) Tailoring - Dimensions of Tailoring Difficulty

	Individual Activity & Aid Complexity Ratings					
Tailoring Activities & Aids	**Very Low** (point value = 1)	**Low** (point value = 2)	**Nominal** (point value = 3)	**High** (point value = 4)	**Very High** (point value = 5)	**Corresponding Points**
Parameter Specification	Zero to 50 parms to be initialized.	51 to 100 parms to be initialized.	101 to 500 parms to be initialized.	501 to 1000 parms to be initialized.	1001 or more parms to be initialized	____
Script Writing	Menu driven, 1 to 5 line scripts, 1 to 5 scripts needed.	Menu driven, 6 to 10 line scripts, 6 to 15 scripts needed.	Hand written, 11 to 25 line scripts, 16 to 30 scripts needed.	Hand written, 26 to 50 line scripts, 31 to 50 scripts needed.	Hand written, 51 or more line scripts, 51 or more scripts needed.	____
I/O Report & GUI Screen Specification & Layout	Automated or standard templates used, 1 to 5 reports/screens needed.	Automated or standard templates used, 6 to 15 reports/screens needed.	Automated or standard templates used, 16 to 25 reports/screens needed.	Hand written or custom designed, 26 to 50 reports/screens needed.	Hand written or custom designed, 51 or more reports/screens needed.	____
Security/Access Protocol Initialization & Set-up	1 security level, 1 to 20 user profiles, 1 input screen/user	2 security levels 21 to 50 user profiles, 2 input screens/user	3 security levels 51 to 75 user profiles, 3 input screens/user	4 security levels 76 to 100 user profiles, 4 input screens/user	5 or more security levels 101 or more user profiles; 5 or more input screens/user	____
Availability of COTS Tailoring Tools	No tools available	N/A	N/A	N/A	Tools are available	____

Total Point Score = _____

UNIVERSITY OF
SOUTHERN CALIFORNIA

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) AXCIP - Integrator Experience with COTS Integration Processes
4) APCON - Integrator Personnel Continuity

COTS Component Drivers

5) ACPMT - COTS Product Maturity
6) ACSEW - COTS Supplier Product Extension Willingness
7) APCPX - COTS Product Interface Complexity
8) ACPPS - COTS Supplier Product Support
9) ACPTD - COTS Supplier Provided Training and Documentation

Application/System Drivers

10) ACREL - 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

UNIVERSITY OF
SOUTHERN CALIFORNIA

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

UNIVERSITY OF
SOUTHERN CALIFORNIA

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