NRO Cost Group
USC Code Counter Brief
Betsy Legg
July 25, 2006

NCG
Cost Group
Cohesive, Consistent & Traceable Cost Estimates

Content

- **Overview of USC code counter**
  - Benefits
  - Languages
  - Output file
- **Overview of Diff tool**
  - Benefits
  - Languages
  - Output files
  - Explain duplicates
- **Overview of Software Data Base (SDB)**
  - Repository
    - USC output files
    - diff output files
    - contractor information
  - Reports
    - Staffing profiles
    - SLOC per Computer Software Configuration Item (CSCI)
    - SLOC per language
    - DR's

USC Code Counter

NRO has adopted USC as its standard and other government agencies have expressed interest in adopting USC as their standard

Benefits of USC Code Counter

- Open text file that needs to be compiled; it is not an executable
  - Security issue is at a minimum
- Free software and is open to the community
- This does not force contractors to change the way they count code, we request the contractor to run the USC Code Counter in addition to their own code counter.
- Many contractors are willing and have run the USC Code Counter without problems.
- Information is at a file level
  - File level information allows us to break out the Software counts into functionality as defined by developers.

Languages

Languages of the USC Code Counter

– Current languages

Ada	Java
Assembly	Jovial
C/C++	Pascal
Cobol	PL1
Fortran	

– Languages worked on by the NCG (and submitted to USC)

PERL MUL -Markup lang. such as HTML/XML
SQL

– Languages to be worked on

scripting languages	.net type languages
*C#	IDL
*JavaScript	

(*USC is working on these counters)

USC Code Counter Output file

- An example of the USC output file c_outfile.dat

Temporary Project Name

Total Lines	Blank Lines	Comments Whole	Comments Embedded	Compiler Direct.	Data Decl.	Exec. Instr.	Logical SLOC	File Type	Module Name
655	57	372	31	7	59	130	196	CODE	/src/opcode_compare.C
299	56	38	2	12	22	139	173	CODE	/AssignmentTool/mpcat.C
282	35	9	0	3	76	106	185	CODE	/PacketTools/addFrames/addFrames.C
803	69	79	133	12	51	335	398	CODE	/list_tf/TF_Lister.C
135	23	28	2	7	1	38	46	CODE	/gui_proc/CACPP_ClientNetIf.C
260	55	57	2	15	5	88	108	CODE	/gui_proc/BarForm.C

Temporary Project Name

The Totals

Total Lines	Blank Lines	Comments Whole	Comments Embedded	Compiler Direct.	Data Decl.	Exec. Instr.	Number of Files	File SLOC	File Type	SLOC Definition
2842960	501102	614256	49077	118411	364681	1244532	11758	1727624	CODE	Physical
2842960	501102	614256	49077	115447	190759	827329	11758	1133535	CODE	Logical
0	0	0	0	0	0	0	0	0	DATA	Physical

Number of files successfully accessed..... 11758 out of 11758

Ratio of Physical to Logical SLOC..... 1.52

Number of files with :

Executable Instructions > 100 = 4873

Data Declarations > 100 = 446

Percentage of Comments to SLOC < 60.0 % = 10675 Ave. Percentage of Comments to Logical SLOC = 58.5

Diff Tool

- Diff two baselines of code
- The Diff Tool is based on the USC Code Counter, developed by NCG
- Numerous contractors have tested the tool and are satisfied with results
- Validated on Microsoft, Sun and SGI based platforms
- The Diff Tool will
  - diff two baselines of code
  - give standard USC output files for each baseline
  - count duplicate files separately
  - give a Complexity output
    - Counts tokens/operations such as +-/* and groups them under Math operators

Diff Tool

Benefits of Diff Tool

– The same as USC code counter - plus

- We now have insight into what has changed between baselines - new, deleted, modified and unmodified lines
  – We have a better understanding of how much code is actually reused.
- We have insight into the Complexity of the code – how many assignments, trig functions, operators, conditionals, loops, log functions, pointer and preprocessing functions.

Diff Tool Languages

Languages of the Diff Tool

– Languages worked on by the NCG and sent to Contractors

ADA	C/C++
JAVA	

– Languages to be worked on

scripting: .cshrc .brorne	VB, VB.net
C#	JavaScript
Cobol	Assembly
PERL	HTML/XML
SQL	Jovial
PL1	Pascal
Fortran (next in the queue)	

Diff Output file

C/C++ SOURCE LINES OF CODE COUNTING PROGRAM
(c) Copyright 1998 - 2000 University of Southern California, CodeCount (TM)

Modified to compare two base lines of code. Results of compare are printed to this file

Temporary Project Name

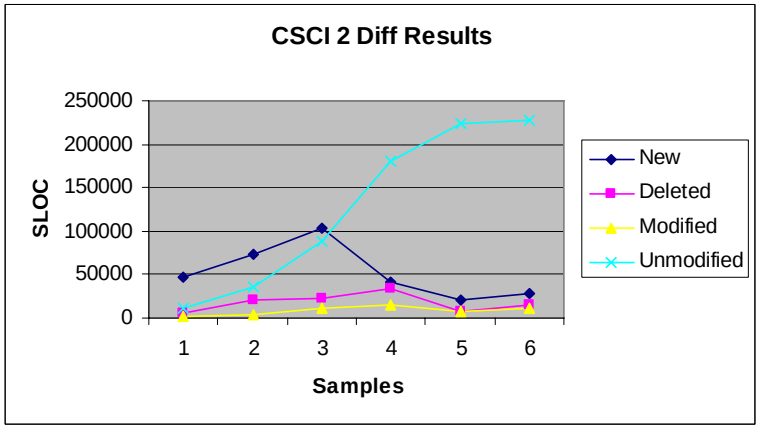
New Lines	Deleted Lines	Modified Lines	Unmodifie Lines	Module Name
0	0	0	196	/src/opcode_compare.C
0	0	0	173	/AssignmentTool/mpcat.C
59	4	51	2698	/src/sc_CompIO.C
0	0	0	46	/src/sc_CompRtn.C
0	22	1	19	/src/sc_CompSactCmdPriParam.C
22	7	7	12	/src/sc_CompSactCmdPubParam.C
6	0	0	12	/src/sc_CompSactDiscrete.C
2	11	11	21	/src/sc_CompSactFltCmnd.C
0	12	5	11	/src/sc_CompSactGndCmnd.C
0	0	0	222	/src/sc_CompSactMeasData.C
13	3	21	63	/src/sc_CompSactUnion.C
0	0	0	66	/src/sc_CompTool.C

.
.
.

Total New Lines	Total Deleted Lines	Total Modified Lines	Total Unmodified Lines
102	59	96	2539

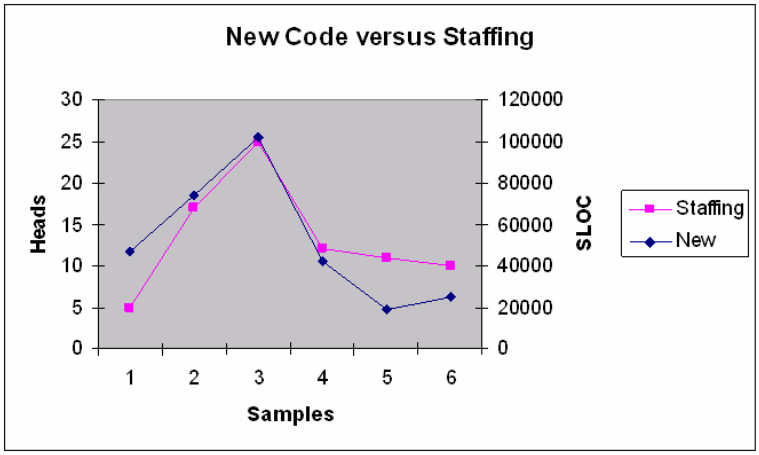
Diff data

- The breakout of the new, deleted modified and unmodified code



New lines versus Staffing

Now we can see a relationship between new lines and staffing.



Diff Complexity Output file

C/C++ SOURCE LINES OF CODE COUNTING PROGRAM
(c) Copyright 1998 - 2000 University of Southern California, CodeCount (TM)

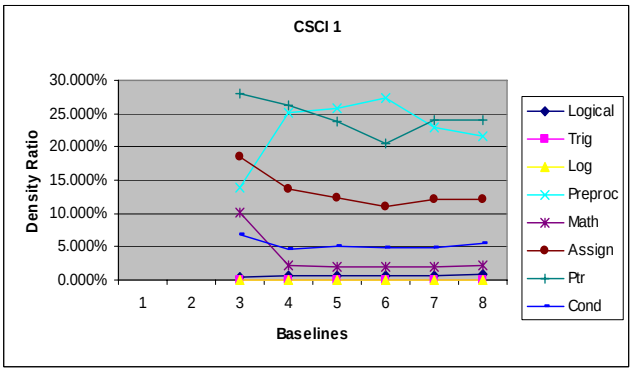
- University of Southern California retains ownership of this copy of software. It is licensed to you. Use, duplication, or sale of this product, except as described in the CodeCount License Agreement, is strictly prohibited. This License and your right to use the software automatically terminate if you fail to comply with any provisions of the License Agreement. Violators may be prosecuted.
This product is licensed to : USC CSE and COCOMO II Affiliates
Note: # stands for baseline; S stands for New, Deleted, Common
Temporary Project Name

#|	S	|Cond	|Logical|	Trig	|	Log	|Preproc|	Math	|Assign|	Ptr	|	Nesting Lvl	|	Filename
A	C	38	2	0	0	16	1	52	15	1,		opcode_compare.C
B	C	38	2	0	0	16	1	52	15	1,		opcode_compare.C
A	C	17	1	0	0	12	2	10	0	4,2,		mpcat.C
B	C	17	1	0	0	12	2	10	0	4,2,		mpcat.C
A	C	14	8	0	0	4	36	66	42	3,1,1,		addFrames.C
B	C	14	8	0	0	4	36	66	42	3,1,1,		addFrames.C
A	C	100	8	0	0	15	55	171	54	23,9,3,		TF_Lister.C
B	C	100	8	0	0	15	55	171	54	23,9,3,		TF_Lister.C
A	C	8	0	0	0	7	0	6	6			CACPP_ClientNetIf.C
B	C	8	0	0	0	7	0	6	6			CACPP_ClientNetIf.C
A	C	1	0	0	0	15	0	23	14			BarForm.C
B	C	1	0	0	0	15	0	23	14			BarForm.C
A	C	1	0	0	0	15	2	10	2	2,		CACPP_BarChartWindow.C
B	C	1	0	0	0	15	2	10	2	2,		CACPP_BarChartWindow.C
A	C	3	0	0	0	14	0	0	3			CACPP_Connect.C
B	C	3	0	0	0	14	0	0	3			CACPP_Connect.C

Complexity Density factor

When you divide the number of operators by the number of lines of code, one can see the density ratio in the code

This could lead to a complexity ratio.



What to do with duplicate files

- We realized early on that duplicate files can occur in a baseline.
- What is a duplicate file?
  - A duplicate file can be a file that is the same code, same file name in multiple places.
  - A duplicate file can be a file with the different code, same filename in multiple places.
- How does the Diff tool handle this issue?
  - Pulls out the list of duplicated and counts then separately and places the results in an output file for duplicates

Software Data Base

- The NCG has on developed a robust Software Database (SDB) that will provide analysts lower level data.
- The SDB was developed to house all the USC Code Count/Diff output files, Complexity and Diff output files, staffing profiles, DR's, cost and hours.
  - Version 1.0 was delivered in September 2005
    - Able to load USC output files into database
  - Version 1.1 was delivered in December 2005
    - Able to map the data in WBS levels
  - Version 2.0 delivered in April 2006
    - Able to get Diff output file loaded and mapped in the Database

Information housed in SDB

The SDB can generate reports from the information imported.

Information imported into the SDB

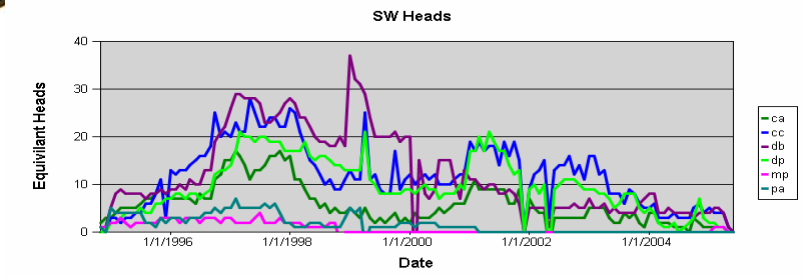
- USC output files
- Diff Tool output files
- Diff Tool Complexity output files
- Diff Tool Diff results output files
- Diff Tool Duplicate output files
- Contractor code counts
- Cost and hours
- DR information
- Staffing profiles

SDB Reports

SDB Reports include

- A report to display CSCIs for a program and contractor
- A report to display Staffing for selected CSCIs
- A report to display DRs per CSCI and Priority for a given Program
- A report to display SLOC for selected WBS or Language
- A report to display SLOC for selected CSCI or WBS element

Example of SDB reports - staffing

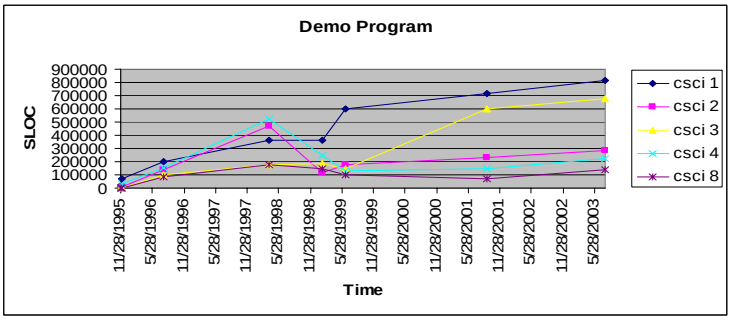


Sample data

Demo Program

	ca	cc	db	dp	mp	pa
11/1/1994	2	0	1	0	1	1
12/1/1994	3	0	1	0	1	0
1/1/1995	3	3	5	2	2	5
2/1/1995	4	3	8	2	2	4
3/1/1995	5	2	9	3	3	4
4/1/1995	5	3	8	4	2	4
5/1/1995	5	3	8	4	1	4
6/1/1995	5	4	8	4	2	4

Example of SDB reports – SLOC per CSCI



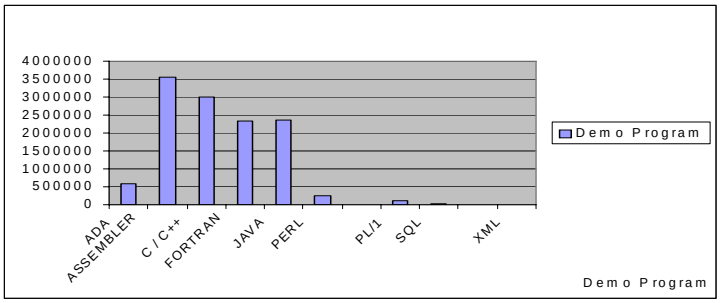
Sample data

Demo Program - SLOC Breakout Detail

	11/28/1995	12/7/1995	7/29/1996	4/2/1998	2/4/1999	6/17/1999	9/13/2001	7/28/2003
csci 1	14494	71406	203112	364732	364732	598308	715943	812237
csci 2	0	4941	135373	465715	118306	177762	228628	284922
csci 3	0	2833	100331	175061	184200	142885	599785	676079
csci 4	0	30425	155158	526255	242822	131365	145516	224810
csci 8	0	0	86777	177761	145745	100966	66275	141565

Example of SDB reports – SLOC per Language

-

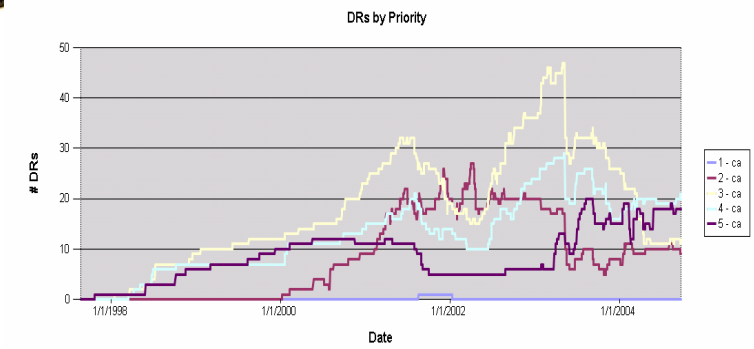


Sample data

SLOC Per Language Detail

	ADA	ASSEMBLER	C / C++	FORTRAN	JAVA	PERL	PL/1	SQL	XML
Demo Program	578493	3569003	2987960	2345634	2353160	239074	122545	37239	1034

Example Data behind the SDB DR Report



Sample data

Demo Program - DRs by Priority

	1 - ca	2 - ca	3 - ca	4 - ca	5 - ca
4/29/2003	0	18	46	28	13
4/30/2003	0	18	47	28	13
5/1/2003	0	17	47	28	13
5/6/2003	0	17	47	28	13
5/7/2003	0	17	47	29	13
5/10/2003	0	17	47	29	13

Bottom Line

- Standardization and Consistency of Code Counting
  - Aerospace and USC are working hard to provide the community counters for all languages
- Continued development of Diff tool for all languages
- Continued development of Software Database

Consistent, Standardized measurements, better tools to analysis data, one central place for data

Backup Slide

- Contact information
  - Sarah Capellari
    - scapellari@mcri.com
    - 703 633 2167
  - Paul Cymerman
    - pcymerman@tecolote.com
    - 703 633 2120
  - Jill Dunn (GPOC)
    - dunnjull@nro.mil
    - 703 633 2141
  - Betsy Legg
    - elegg@tecolote.com
    - 703 633 2102