“Telling More“: a 2006 NASA SARP project (Extended Executive Briefing)

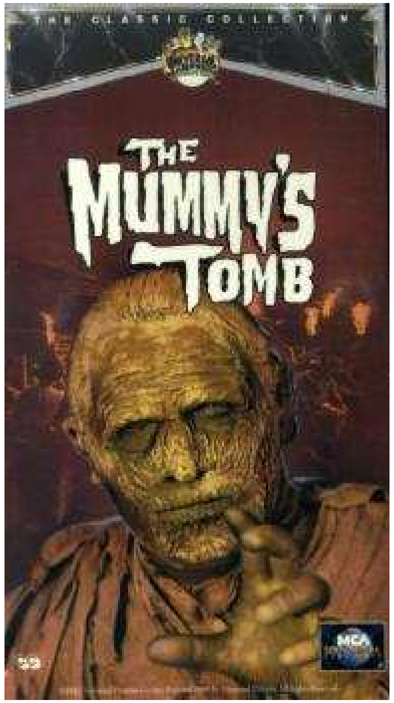
Tim Menzies
Lane Department of Computer Science and Electrical Engineering
West Virginia University
tim@menzies.us

May 30, 2006

SOUND BITES

- Use it or Lose it
- Empirical knowledge factories
- The only good answers destroy the question.
- What's up with SILAP?
- Face the hardest, but don't forget the aveage.

Problem: Use it or Lose it



- NASA data = active repository?
- Or a data tomb?
  - write once;
  - read never;
  - buried;
  - doomed;

- NASA IV&V is in a unique position to review and comment on the entire NASA software enterprise.
  - We *see more*, but what have we *learned*?
  - What can we *tell*?

Approach: use AI

1. Apply AI machinery to NASA data repositories.
   - When data is plentiful, use *data miners*.
     (e.g. cost and defect estimation)
   - When data is scarce and domain intuitions exist:
     - build a *what-if* simulator for those intuitions,
     - Monte Carlo the simulator,
     - goto 1

     (e.g. this briefing)
   - When data and intutions exist, use bayesian belief nets;
     - e.g. Dabney, Fenton, etc etc
   - When complex domain models exist, use semantic web tools to generalize from *here* to *there*.

- Find and collect avaialble data
  - Maintaining security and confidentiality requirements.
- Match data sources to available machines
- Apply the machines.

Importance & Benefits

If an organizations spends millions of data collection and archiving...

- It should spend tens of thousands (or more) analyzing that data.

Surprises, found before by SARP projects:

- Largest source of post-launch *deep space* anomalies is *ground systems* [Lutz,2004]
- Common conflation of *severity* and *priority* in NASA defect logs [me,2002]
- *Small* changes in data $\Rightarrow$ *massive* changes in cost estimation [me,Hihn,2005]
- Static code measures surprisingly good and predict for issues [me, 2005]

And the discoveries continue:

- IV&V tasks often the *same*, despite processing different *projects* [me, today]

What might we learn tomorrow?

Accomplishments

Example

Learning must be adjusted to fit the available data.

Coming soon to an IV&V center near you:

- Bayesian belief networks for monitoring + diagnosis of IV&V processes;
- The "LINKER" database describing cost/effort data for IV&V tasks;
- The CRADLE tool storing Constellation artifacts;
- etc

For now, what we can access are:

- The current IV&V work breakdown structure (WBS)
- SILAP: the IV&V model that selects WBS tasks
- MB1: 433 NASA sub-systems (CSCIs) expressed using the SILAP variables;
- Four subsets of MB1:
  - Scored 1,2,3,4 for "distance" to earth
  - MB1.gs1 - ?ground systems : 20 records
  - MB1.eo2 - ? earth orbit : 265 records
  - MB1.xf3 - ?? : 93 records
  - MB1.go4 - most distant : 57 records
- Important note: sanitized data (no project names), and oly available to researchers who apply to NASA IV&V.

So, what can we "tell more" about SILAP?

SILAP


```
function CO1(   tmp) { // Consequence
    // HS2 =  Human Safety; AS2 =  Asset Safety; PF2 =  Performance
    tmp = 0.35*value("AS2") + 0.65 *value("PF2")
    return (value("HS2") > tmp) ? value("HS2") : tmp;
}
```

```
function EP1()  { // Error Potential
    //      Development + Process      + Software characteristics
    return  0.579*DV2() + 0.249*PR2()  + 0.172*SC2() }
```

```
function SC2() { // Software Characteristic
    //              Complexity + Degree of Innovation  + Size of system
    return  0.547*value("CX3") + 0.351*value("DI3")    + 0.102*value("SS3") }

function DV2()  { // Development
    //       Experience        + development organization
    return  0.828*value("EX3") + 0.172*value("DO3") }

function PR2()  { // Performance
    //        Re-use approach + artifact maturity  + formality
    return 0.226*value("RA3") + 0.242*value("AM3") + formality() }
```

```
function formality() {
    //     Use of standards    + Use of CM           + CMM level
    return 0.0955*value("US3") + 0.0962*value("UC3")  + 0.0764*value("CL3")
    //Use  formal reviews + use defect tracking + use  risk management tool
     +0.1119*value("FR3") + 0.0873*value("DT3") + 0.0647*value("RM3") }
```

Questions about SILAP

- Q1: is SILAP worthy of study?
  - A1a: Yes; it is how the Fairmont IV&V civil servants select IV&V activities
  - A1b: Heck yes: the (growing) experience base at Fairmont is based around SILAP
  - A1c: The civil servants at IV&V are to be commended for offering an explicit, publically accessible, model of their work practices.
    - This record lets outsiders (like me) review and offer iprovements to their work practices.
    - How many other organizations can say they have done the same?
- Q2: does it have the right variables?
  - ?
- Q3: does it have the right connections between variables?
- Q4: do the connections have the right weights?
  - 0.0647, 0.0764, 0.0873, 0.0955, 0.0962, 0.102, 0.1119, 0.172, 0.172, 0.226, 0.242, 0.249, 0.35, 0.351, 0.547, 0.579, 0.65, 0.828

Our answer to Q3 and Q4 will be “unimportant, for now; need to fix something else first”.

The rest of this talk: what is the something else.

SLIAP selects tasks from the IV&V work breakdown structure

Introduction
Problem
Approach
Importance & Benefits
Accomplishments
❖ Example
❖ SILAP
❖ Questions about SILAP
❖ WBS
❖ WBS (cont.)
❖ SILAP inputs
❖ SILAP inputs (cont.)
❖ Selected WBS tasks
❖ Now, the fun begins
❖ Ep/Con
❖ Rx + Treatment
Relevance to NASA
Next steps

wbs	factor	CO11	CO12	CO13	CO14	CO15	EP11	EP12	EP13	EP14	EP15
1.1	Management and Planning of IV&V	X	X	X	X	X	X	X	X	X	X
1.2	Issue and Risk Tracking		X	X	X	X		X	X	X	X
1.3	Final Report Generation		X	X	X	X		X	X	X	X
1.4	IV&V Tool Support		X	X	X	X		X	X	X	X
1.5	Management & Technical Review Support	X	X	X	X	X	X	X	X	X	X
1.6	Criticality Analysis	X	X	X	X	X	X	X	X	X	X
2.1	Reuse Analysis			Y	Y	Y					
2.2	Software Architecture Assessment			X	X	X					
2.3	System Requirements Review			X	X	X				X	X
2.4	Concept Document Evaluation				Z	Z					Z
2.5	SW/User Requirements Allocation Analysis				Z	Z					Z
2.6	Traceability Analysis				Z	Z					Z
3.1	Traceability Analysis - Requirements		X	X	X	X				X	X
3.2	Software Requirements Evaluation			X	X	X				X	X
3.3	Interface Analysis - Requirements				X	X			X	X	X
3.4	System Test Plan Analysis			X	X	X					
3.5	Acceptance Test Plan Analysis					X					
3.6	Timing and Sizing Analysis									Z	Z

X = "or";
Y = for $RA3 > 2$;
Z = for human-rated flights

CO1 = "consequences of failure"
EP1 = "error potential"
$\{$CO1, EP1$\}$ computed via SILAP

Work breakdown structure (cont.)

Introduction

Problem

Approach

Importance & Benefits

Accomplishments

- ❖ Example
- ❖ SILAP
- ❖ Questions about SILAP
- ❖ WBS
- ❖ WBS (cont.)
- ❖ SILAP inputs
- ❖ SILAP inputs (cont.)
- ❖ Selected WBS tasks
- ❖ Now, the fun begins
- ❖ Ep/Con
- ❖ Rx + Treatment

Relevance to NASA

Next steps

wbs	factor	CO11	CO12	CO13	CO14	CO15	EP11	EP12	EP13	EP14	EP15
4.1	Traceability Analysis – Design							X	X	X	X
4.2	Software Design Evaluation				X	X				X	X
4.3	Interface Analysis – Design					X				X	X
4.4	Software FQT Plan Analysis		X	X	X	X					
4.5	Software Integration Test Plan Analysis									X	X
4.6	Database Analysis								X	X	X
4.7	Component Test Plan Analysis										X
5.1	Traceability Analysis – Code					X		X	X	X	X
5.2	Source Code and Documentation Evaluation				X	X			X	X	X
5.3	Interface Analysis – Code				X	X			X	X	X
5.4	System Test Case Analysis				X	X					
5.5	Software FQT Case Analysis				X	X					
5.6	SW Integration Test Case Analysis										X
5.7	Acceptance Test Case Analysis					X					
5.8	SW Integration Test Procedure Analysis										X
5.9	SW Integration Test Results Analysis									X	X
5.1	Component Test Case Analysis										X
5.11	System Test Procedure Analysis					Z					
5.12	Software FQT Procedure Analysis					Z					
6.1	Traceability Analysis – Test		X	X	X	X					X
6.2	Regression Test Analysis									Z	Z
6.3	Simulation Analysis					Z					
6.4	System Test Results Analysis				X	X					
6.5	Software FQT Results Analysis				X	X					
7.1	Operating Procedure Evaluation					Z					
7.2	Anomaly Evaluation					Z					
7.3	Migration Assessment					Z					
7.4	Retirement Assessment					Z					

SILAP input parameters: distributions

433 records (all)

HS2	3 =0.00 5=0.02 2=0.02 4=0.05 1=0.35 0=0.56
AS2	5 =0.07 2=0.07 4=0.10 3=0.10 1=0.67
PF2	2 =0.09 5=0.13 4=0.16 1=0.19 3=0.43
EX3	5 =0.05 4=0.14 2=0.17 3=0.21 1=0.44
DO3	0.5=0.01 3=0.02 1=0.08 5=0.11 2=0.33 4=0.45
US3	5 =0.04 3=0.21 2=0.24 1=0.50
UC3	3 =0.07 2=0.28 1=0.65
CL3	2 =0.00 5=0.04 4=0.20 3=0.76
FR3	3 =0.01 2=0.34 1=0.65
DT3	4 =0.01 3=0.31 2=0.34 1=0.35
RM3	1 =0.18 3=0.36 2=0.45
RA3	2 =0.07 3=0.11 4=0.12 5=0.12 1=0.58
AM3	3 =0.15 2=0.21 1=0.64
CX3	5 =0.04 2=0.11 4=0.12 1=0.31 3=0.41
DI3	4 =0.02 3=0.02 2=0.12 1=0.85
SS3	1 =0.06 4=0.07 2=0.30 3=0.57

How to read these tables

- e.g. the last line of the above;
- sorted left-to-right as rarest-to-more-frequent;
- Least common SS3 value = “1”: occurs 6% of the time;
- Most common SS3 value = “3”: occurs 57% of the time;
- “2” occurs 30% of the time;

265 records (eo2)

HS2	3 =0.01 5=0.01 2=0.02 4=0.08 1=0.41 0=0.48
AS2	5 =0.03 2=0.08 4=0.09 3=0.11 1=0.69
PF2	5 =0.09 2=0.10 4=0.19 1=0.25 3=0.37
EX3	5 =0.07 2=0.15 3=0.17 4=0.17 1=0.44
DO3	0.5=0.02 3=0.03 5=0.04 4=0.45 2=0.46
US3	5 =0.07 3=0.14 2=0.22 1=0.57
UC3	3 =0.03 2=0.16 1=0.81
CL3	2 =0.00 5=0.07 4=0.14 3=0.79
FR3	3 =0.00 2=0.09 1=0.91
DT3	4 =0.01 3=0.11 2=0.31 1=0.56
RM3	1 =0.16 2=0.30 3=0.54
RA3	3 =0.00 4=0.01 2=0.12 5=0.12 1=0.74
AM3	3 =0.19 2=0.20 1=0.61
CX3	5 =0.04 2=0.07 4=0.11 3=0.37 1=0.41
DI3	4 =0.01 3=0.01 2=0.08 1=0.91
SS3	1 =0.01 4=0.04 2=0.41 3=0.55

93 records (xf3)

HS2	2=0.02 1=0.10 0=0.88
AS2	5=0.03 2=0.07 3=0.09 4=0.14 1=0.67
PF2	2=0.08 1=0.10 5=0.11 4=0.15 3=0.57
EX3	4=0.12 2=0.14 3=0.24 1=0.50
DO3	1=0.18 5=0.37 4=0.45
US3	3=0.21 2=0.24 1=0.55
UC3	3=0.03 1=0.36 2=0.61
CL3	4=0.39 3=0.61
FR3	1=0.29 2=0.71
DT3	1=0.03 2=0.48 3=0.49
RM3	1=0.23 2=0.77
RA3	3=0.12 1=0.39 4=0.49
AM3	3=0.03 2=0.13 1=0.84
CX3	5=0.03 4=0.11 2=0.20 1=0.21 3=0.46
DI3	3=0.03 2=0.13 1=0.84
SS3	2=0.13 1=0.27 3=0.60

SILAP input parameters: distributions (cont.)

57 records (go4)

HS2	0=0.43 1=0.57
AS2	2=0.04 4=0.04 3=0.11 5=0.27 1=0.55
PF2	4=0.09 2=0.09 1=0.11 5=0.29 3=0.43
EX3	5=0.02 4=0.04 3=0.09 2=0.41 1=0.45
DO3	5=0.05 1=0.30 4=0.64
US3	1=0.27 3=0.27 2=0.46
UC3	1=0.27 3=0.32 2=0.41
CL3	4=0.23 3=0.77
FR3	3=0.07 1=0.25 2=0.68
DT3	3=1.00
RM3	1=0.25 3=0.27 2=0.48
RA3	4=0.05 1=0.30 3=0.64
AM3	3=0.21 1=0.38 2=0.41
CX3	1=0.09 5=0.11 2=0.12 4=0.20 3=0.48
DI3	3=0.05 4=0.09 2=0.30 1=0.55
SS3	2=0.18 3=0.82

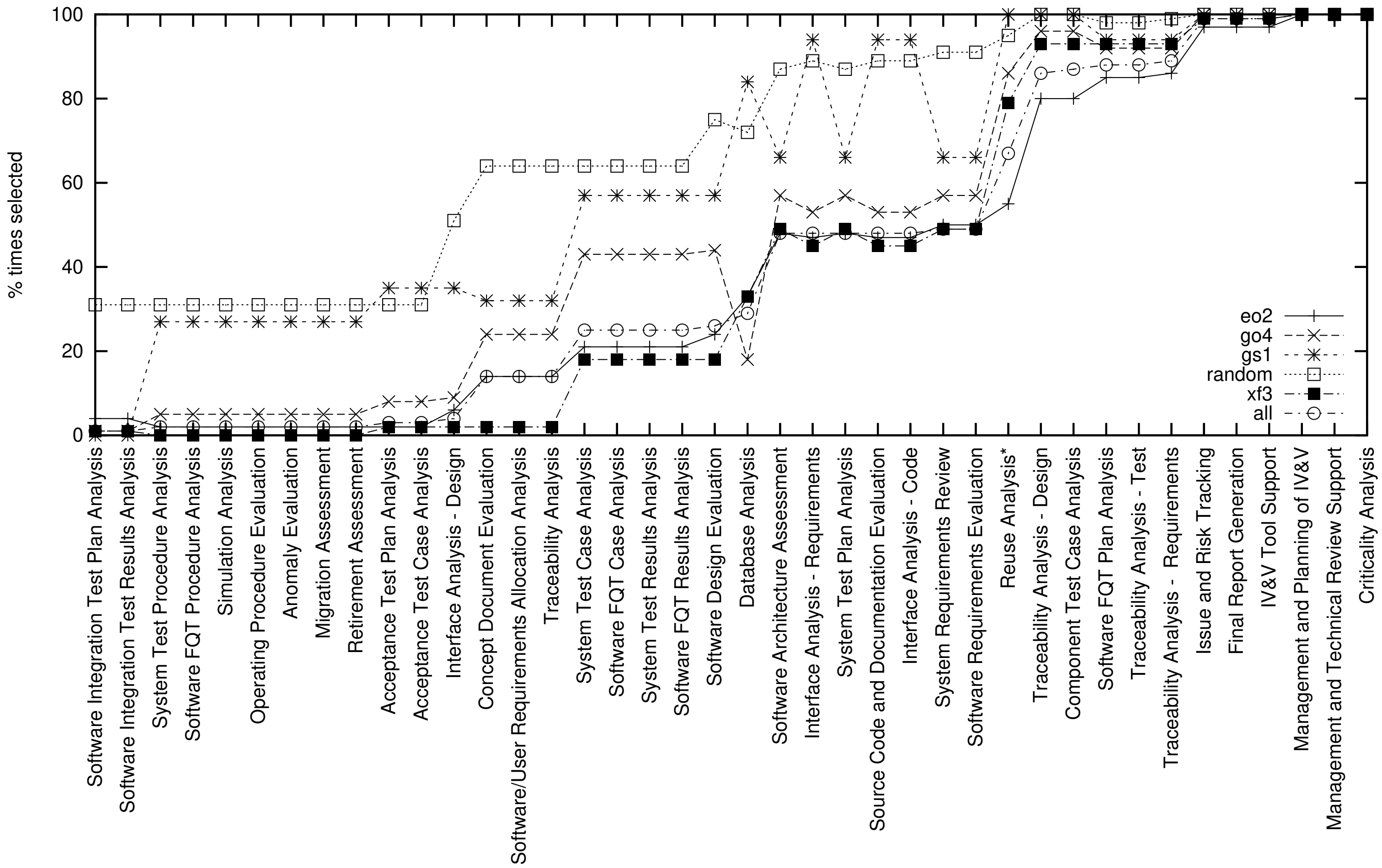
20 records (gs1)

HS2	4=0.05 2=0.05 5=0.26 0=0.63
AS2	5=0.16 4=0.21 1=0.63
PF2	4=0.05 2=0.05 1=0.11 5=0.26 3=0.53
CX3	4=0.11 1=0.16 2=0.32 3=0.42
EX3	3=1.00
DO3	2=1.00
US3	3=1.00
UC3	1=1.00
CL3	3=1.00
FR3	2=1.00
DT3	2=1.00
RM3	2=1.00
RA3	5=1.00
AM3	1=1.00
DI3	1=1.00
SS3	4=1.00

Comment: certain minority project types are radically different to the majority.

Selected WBS tasks

500 times * {pick SILAP inputs from known distributions; compute (EP1,CO1); select relevant WBS tasks}



xf3m eo2 have very similar patterns.

go4 nearly has the same

Size of $xf3 + eo2 + go4$ is 265 + 93 + 57 = $\frac{415}{433} = 96\%$ of the data.

So current version of SILAP curiously insensitive to most known project variations.

Say what???

Now, the fun begins

"The way to do research is to attack the facts at the point of greatest astonishment."
- Celia Green

"The only good answers destroy the question."
- Susan Sontag

"In any field, find the strangest thing and then explore it."
- John Archibald Wheeler

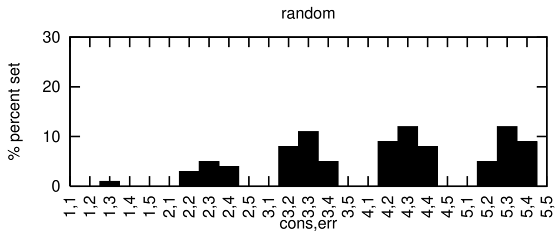
"What the heck is going on with SILAP?"
- NASA Fairmont process improvement group

Error Potential/ Consequence Distributions

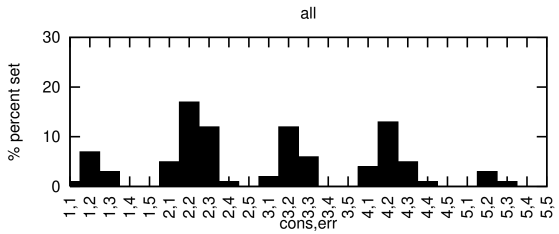
Introduction
Problem
Approach
Importance & Benefits
Accomplishments
❖ Example
❖ SILAP
❖ Questions about SILAP
❖ WBS
❖ WBS (cont.)
❖ SILAP inputs
❖ SILAP inputs (cont.)
❖ Selected WBS tasks
❖ Now, the fun begins
❖ Ep/Con
❖ Rx + Treatment
Relevance to NASA
Next steps

Log the (CO1,EP1) values calculated by SILAP for different MB1 subsets

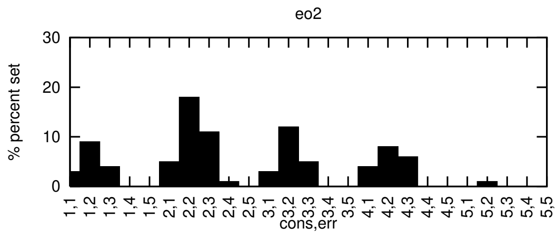
500 randomly generated records:



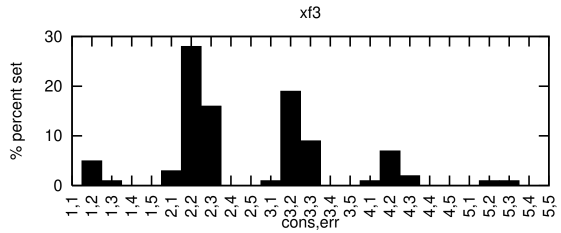
433 records:



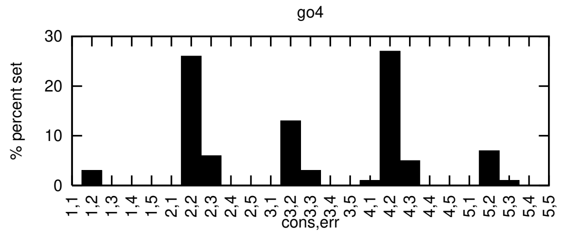
265 recods:



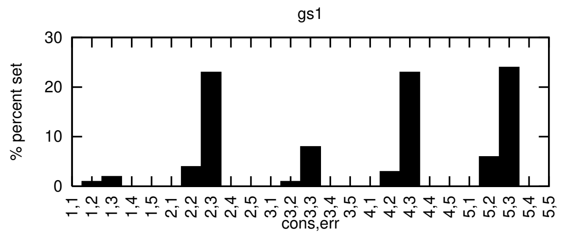
93 records:



57 records:



20 records:



Comment:

- in the majority of cases, the MB1 data set does *not* select for $CO1 \geq 4$ or $EP1 \geq 4$

Diagnosis and Treatment

- There is a bias in the WBS task selection:
  - Most tasks need $CO1 \geq 4$ or $EP1 \geq 4$.
- Diagnosis: the current SILAP/WBS selection tool differentiates *hard problems*;
  - But task selection is very insenstive to *ordinary problems* $CO1 < 4$ or $EP1 < 4$.
- So, treatment option one:
  - Applied one-size-fits-all WBS tasks to “ordinary” projects;
  - Reserve SILAP for “hard” projects.
- Treatment option two: modify SILAP/WBS
  - Take the time to explore “ordinary projects”
  - How do they differ? What should we do about them?
- Option two is preferred....
  - If it is desired to optimize the day-to-day running of routine IV&V (which, according to MB1, is the norm)

Relevance to NASA

- All the above where from NASA data sets;

Next steps

- Always eager to process data from new sources.