SOFTWARE SIZE ESTIMATING
INDUSTRY PERCEPTION

"WE CAN'T ESTIMATE SIZE"
PERCEIVED PROBLEMS/OBSTACLES TO ACCURATE SOFTWARE SIZING

* REQUIREMENTS VOLATILITY
* SIMPLE METHODS – NO RIGOR
* RETRO FITTING CODE ESTIMATES TO POLITICAL EXPECTATIONS
* POINT ESTIMATES AND RISK PADDING
* CONSISTENT DEFINITIONS AND COUNTING PROCEDURES
* LIMITED METHODS
* ABSENSE OF HISTORIC DATA
* TECHNICAL vs. NON TECHNICAL COMMUNICATION
QUANTITATIVE METHODS TO

* BOUND SIZE
* DETERMINE UNCERTAINTY
* IDENTIFY RISK
SIZE PLANNER
(BASED ON MULTIPLE APPROACHES)

BENEFITS:

* View problem from different perspectives
* Use statistical techniques to combine answers
* Convergence/divergence confirm quality of estimate
* Refine as new information becomes available
SIZE PLANNER CONCEPT

- Fuzzy Logic
- Function Points
- Standard Components
- New Reused Retested

BAYESIAN STATISTICAL COMBINATION PROCESS

3 POINT SIZE ESTIMATE

COPYRIGHT DSM (C) ALL RIGHTS RESERVED
SIZE PLANNER
STATISTICAL METHODS
BAYSIAN WEIGHTING

* BAYSIAN WEIGHTING IS A AVERAGING TECHNIQUE

* IT GIVES MORE WEIGHT TO THE EXPECTED VALUES THAT HAVE THE LEAST AMOUNT OF UNCERTAINTY

* THE UNCERTAINTY IS MEASURED BY THE STANDARD DEVIATION
BAYSIAN WEIGHTING
(CONT.)

* IN SIZE PLANNER IT IS USED WITHIN EACH TECHNIQUE AND TO COMBINE THE DIFFERENT TECHNIQUES

* RESULTING ESTIMATES WILL HAVE A REDUCED UNCERTAINTY
EXPONENTIAL SMOOTHING

* EXPONENTIAL SMOOTHING IS A UPDATE CONVERGENCE TECHNIQUE

* IN SIZE PLANNER THIS TECHNIQUE IS USED TO PICK UP GROWTH OR REDUCTION TRENDS AND UPDATE THE ESTIMATE TO REFLECT THOSE TRENDS

* IF SIGNIFICANT CHANGES TAKE PLACE FROM PREVIOUS ESTIMATE TO CURRENT ESTIMATE, THE EXPONENTIAL SMOOTHING MAY NOT BE SENSITIVE ENOUGH TO COMPENSATE. IN THIS CASE THE ESTIMATE SHOULD BE STORED AS A NEW FILE
FUZZY LOGIC SIZING
FUZZY LOGIC SIZING

SIZE DISTRIBUTION ESLOC * 1000
FUZZY LOGIC SIZING

ESTIMATOR PROVIDES:
* APPLICATION TYPE SELECTION
* OVERALL SIZE CATEGORY SELECTION
* REFINED SIZE RANGE SELECTION

QSM DATA BASE PROVIDES:
* SIZE STATISTICAL PROFILES BY APPLICATION
* STATISTICAL PROFILES BY SIZE CATEGORY

BAYSIAN STATISTICS PROVIDES:
* METHOD TO COMBINE AN WEIGH INTUITIVE SELECTIONS WITH HISTORICAL DATA

COPYRIGHT QSM (C) ALL RIGHTS RESERVED
### Fuzzy Logic

<table>
<thead>
<tr>
<th>% Prob of Not Exceeding Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1 Std Dev</td>
</tr>
<tr>
<td>Exp</td>
</tr>
<tr>
<td>+1 Std Dev</td>
</tr>
</tbody>
</table>

#### Graph

- **CURRENT EST**

- **X**: CURRENT EST
- **Y**: % Prob of Not Exceeding Size

<table>
<thead>
<tr>
<th>CURRENT EST</th>
<th>% Prob of Not Exceeding Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>-18</td>
</tr>
<tr>
<td>8</td>
<td>-16</td>
</tr>
<tr>
<td>10</td>
<td>-14</td>
</tr>
<tr>
<td>99</td>
<td>+1</td>
</tr>
</tbody>
</table>

The graph shows the probability of a value not exceeding a certain size given different standard deviations.
STATISTICAL METHODS

TWO PRIMARY TECHNIQUES:

1. BAYSIAN WEIGHTING
2. EXPONENTIAL SMOOTHING
FUZZY LOGIC SIZING

STRENGTHS:

* Requires modest input
* Can be applied very early
* Probability distribution bounds range

WEAKNESSES:

* Not highly precise
* Intuition - subjective
FUNCTION POINT SIZING
FUNCTION POINT SIZING

TECHNIQUE:

* IDENTIFY THE NUMBER OF FUNCTION PROVIDING ELEMENTS IN EACH OF 5 CATEGORIES

* CATAGORIZE BY PROCESSING COMPLEXITY

* CALCULATE NUMBER OF FUNCTION POINTS

COPYRIGHT DSM (C) ALL RIGHTS RESERVED
FUNCTION POINT SIZING

TECHNIQUE: (CONTINUED)

* IDENTIFY PREDOMINANT LANGUAGE

* CONVERT FUNCTION POINTS TO SLOC
**FUNCTION POINT REPORT**

Date: 09-25-1987
Time: 16:46:31

**User External Inputs**

- **Simple**: 0
- **Moderate**: 0
- **Average**: 10
- **Complex**: 2
- **Highly Complex**: 0

**Total**: 12

**User External Outputs**

- **Simple**: 0
- **Moderate**: 0
- **Average**: 8
- **Complex**: 3
- **Highly Complex**: 0

**Total**: 13

**User External Inquiries**

- **Simple**: 0
- **Moderate**: 0
- **Average**: 15
- **Complex**: 0
- **Highly Complex**: 0

**Total**: 15

**Logical Master Files Interfaces**

- **Simple**: 0
- **Moderate**: 0
- **Average**: 8
- **Complex**: 0
- **Highly Complex**: 0

**Total**: 8

**Total Expected Function Points**: 50
**Total Lines of Code**: 27084
**Total Weighted Function Points**: 270
**Standard Deviation**: 5299
STANDARD COMPONENT SIZING
FUNCTION POINT SIZING

STRENGTHS:
* LANGUAGE INDEPENDENT
* ENHANCES COMMUNICATION WITH USERS/NON TECHNICAL PEOPLE

WEAKNESSES:
* MUST BE TO LOW LEVEL DESIGN TO APPLY WITH ACCURACY
* AMBIGUOUS DEFINITION OF PARAMETERS
* SUBJECTIVE COMPLEXITY ADJUSTMENTS
* LANGUAGE EXPANSION FACTORS UNPROVEN
* EVOLVING TECHNIQUE - ALGORITHM ADJUSTMENT
STANDARD COMPONENT SIZING

TECHNIQUE APPLICATION:

* FOR NEWLY DEVELOPED SOFTWARE
* USED FROM CONCEPT THROUGH SYSTEM TEST
* FOR ALL TYPES OF SYSTEMS
STANDARD COMPONENT SIZING

SYSTEM DECOMPOSITION

ABstraction Level

SUBSYSTEMS

MODULeS
SCREENS
REPORTS

FILES
PROGRAMS
SLOC
BYTES
STANDARD COMPONENT SIZING

TOTAL OF 12 COMPONENT INPUTS:

* SUBSYSTEMS
* MODULES
* SCREENS
* REPORTS
* INTERACTIVE PROGRAMS
* BATCH PROGRAMS
* SLOC
* FILES
* BYTES
* BITS
* WORDS
* OBJECT INSTRUCTIONS
STANDARD COMPONENT SIZING
CALIBRATION:

* PROVIDES ABILITY TO TUNE SIZE PLANNER TO YOU OWN HISTORY

* HISTORICAL DATA COLLECTION FACILITY IS BUILT INTO SYSTEM

* ESTABLISHES RELATIONSHIPS BETWEEN A COMPONENT AND SLOC
STANDARD COMPONENT SIZING

SIZING HISTORY ANALYSIS REPORTS:

* Show average number of components per system

* Variability associated with the average

* Ranking of components predictive value

* Establish statistical ratio relationships between all of the standard components

COPYRIGHT DSM (C) ALL RIGHTS RESERVED
STANDARD COMPONENT SIZING

ESTIMATOR PROVIDES:

* 3 POINT ESTIMATES LOW, MOST LIKELY, HIGH FOR ALL COMPONENTS APPLICABLE AT THAT TIME.

* A CONFIDENCE RATING, LOW, MODERATE, HIGH FOR EACH COMPONENT ESTIMATE. THIS IS USED AS A CHECK AND DISCRIMINATOR ON CONFLICTING ESTIMATED VALUES.
STANDARD COMPONENT SIZING

ESTIMATES:

* EACH COMPONENT IS TURNED INTO A SLOC ESTIMATE

* EACH SLOC ESTIMATE HAS A MEASURED STANDARD DEVIATION DETERMINED BY THE 3 POINT ESTIMATE

* BAYSIAN STATISTICAL WEIGHTS ARE USED TO COMBINE COMPONENT ESTIMATES
STANDARD COMPONENT SIZING

STRENGTHS:

* USED THROUGHOUT LIFECYCLE

* CAN BE CUSTOMIZED THROUGH YOUR OWN DATA

* USES 3 POINT RANGES TO QUANTIFY UNCERTAINTY
## STANDARD COMPONENT SIZING REPORT

**Date:** 09-27-1987  
**Time:** 11:21:17

---

Project: RTD  
Language: Cobol

<table>
<thead>
<tr>
<th>Component</th>
<th>Low</th>
<th>Most Likely</th>
<th>High</th>
<th>Conf.</th>
<th>Weight to QSM Data</th>
<th>Expected SLOC</th>
<th>Std Dev SLOC</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLOC</td>
<td>15000</td>
<td>25000</td>
<td>35000</td>
<td>2</td>
<td>82</td>
<td>25000</td>
<td>5000</td>
</tr>
<tr>
<td>Obj Instr.</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Bits</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>82</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Bytes</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>32</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Words</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>82</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Files</td>
<td>6</td>
<td>10</td>
<td>12</td>
<td>1</td>
<td>10</td>
<td>20314</td>
<td>3152</td>
</tr>
<tr>
<td>Modules</td>
<td>25</td>
<td>35</td>
<td>45</td>
<td>2</td>
<td>10</td>
<td>20197</td>
<td>2885</td>
</tr>
<tr>
<td>Subsystems</td>
<td>8</td>
<td>9</td>
<td>12</td>
<td>3</td>
<td>10</td>
<td>21186</td>
<td>2270</td>
</tr>
<tr>
<td>Screens</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>82</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Reports</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>94</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Interactive</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Batch Pgms</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>82</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Combined Weighted Solution:** 21238 ± 1491

---

PgUp-Prior Page  PgDn-Next Page  Esc-Exit
STD COMPONENTS ESTIMATES

% PROB OF NOT EXCEEDING SIZE

-1 STD DEV  EXP  +1 STD DEV

SLOC

1000

15 17 19 21 23 25 27

CURRENT EST
NEW, MODIFIED, REUSED SIZING
NEW, MODIFIED, REUSED SIZING

HISTORY: (CONT.)

* TAUSWORTHE IDENTIFIED FOLLOWING TYPES OF CODE

1. NEW CODE
2. REUSED CODE
   2A. ADDED
   2B. CHANGED
   2C. DELETED (LINE BY LINE)
   2D. REMOVED (WHOLE BLOCKS)
   2E. RETESTED
NEW, MODIFIED, REUSED SIZING

TECHNIQUE:

* RUN FUZZY LOGIC, STANDARD COMPONENT AND FUNCTION POINTS TO ESTIMATE NEW SLOC

* SIZE PLANNER AUTOMATICALLY POSTS 3 POINT RANGE RESULTS IN NEW ROW

* FILL IN 3 POINT ESTIMATES FOR EACH ADDITIONAL CATEGORY

* USE 1%, 50% AND 99% SIZE PROBABILITY VALUES AS RANGE ESTIMATES IN SLIM
## New, Reused, and Rebuilt

<table>
<thead>
<tr>
<th>Lines of Code</th>
<th>Low</th>
<th>Most Likely</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEW</td>
<td>_16765</td>
<td>_21238</td>
<td>_25711</td>
</tr>
<tr>
<td>REUSED</td>
<td>_25000</td>
<td>_35000</td>
<td>_45000</td>
</tr>
<tr>
<td>ADDED</td>
<td>_1000</td>
<td>_3000</td>
<td>_4000</td>
</tr>
<tr>
<td>CHANGED</td>
<td>_2000</td>
<td>_4000</td>
<td>_6000</td>
</tr>
<tr>
<td>DELETED</td>
<td>_250</td>
<td>_450</td>
<td>_650</td>
</tr>
<tr>
<td>REMOVED</td>
<td>_200</td>
<td>_400</td>
<td>_500</td>
</tr>
<tr>
<td>TESTED</td>
<td>_10000</td>
<td>_12000</td>
<td>14000</td>
</tr>
</tbody>
</table>

Enter the absolute highest number of source statements that exist and require no modifications but do require testing with new and modified software.

F1-HELP     F2-RESTORE DEFAULTS     F10-INPUTS COMPLETE
### REBUILT SOFTWARE SIZING REPORT

**Date:** 09-27-1987  **Time:** 11:26:18  

<table>
<thead>
<tr>
<th></th>
<th>Low</th>
<th>Most Likely</th>
<th>High</th>
<th>Expected</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>New</td>
<td>16765</td>
<td>21238</td>
<td>25711</td>
<td>21238</td>
<td>1491</td>
</tr>
<tr>
<td>Reused</td>
<td>25000</td>
<td>35000</td>
<td>45000</td>
<td>35000</td>
<td>3333</td>
</tr>
<tr>
<td>Added</td>
<td>1000</td>
<td>3000</td>
<td>4000</td>
<td>2833</td>
<td>500</td>
</tr>
<tr>
<td>Changed</td>
<td>2000</td>
<td>4000</td>
<td>6000</td>
<td>4000</td>
<td>666</td>
</tr>
<tr>
<td>Deleted</td>
<td>250</td>
<td>450</td>
<td>650</td>
<td>450</td>
<td>66</td>
</tr>
<tr>
<td>Removed</td>
<td>200</td>
<td>400</td>
<td>500</td>
<td>383</td>
<td>50</td>
</tr>
<tr>
<td>Tested</td>
<td>10000</td>
<td>12000</td>
<td>14000</td>
<td>12000</td>
<td>666</td>
</tr>
</tbody>
</table>

**Effective Source Lines of Code:**  

36364  2362
REBUILT, MODIFIED SYSTEMS
% PROB OF NOT EXCEEDING SIZE
-1 STD DEV EXP +1 STD DEV

SLOC

1000

1

10

50

90

99

25

30

35

40

45

CURRENT EST
SIZE PLANNER
RISK PROFILES
SIZE PLANNER RISK PROFILES

FUZZY  STD COMP  FPOINTS  NEW-REUSED

COPYRIGHT QSM, INC. 1986
SIZE PLANNER RISK PROFILES

16% 34% 34% 16%
55K 75K 95K

75% probability the size will not exceed 87K SLOC

COPYRIGHT QSM, INC. 1986
COMBINED CURRENT ESTIMATES
\% PROB OF NOT EXCEEDING SIZE

-1 STD DEV  EXP  +1 STD DEV

SLOC \times 1000

STEEP SLOPE
MORE UNCERTAINTY

STD C  REUSD  WGT
COMBINED CURRENT ESTIMATES
% PROB OF NOT EXCEEDING SIZE

-1 STD DEV  EXP  +1 STD DEV

LARGE DISPERSION AT EXPECTED VALUE INDICATES HIGH UNCERTAINTY IN ANSWER

--- Fuzzy ---- F Pts ---- STD C ---- REUSD --- WGT
CONCEPT

DESIGN

CODE UNITS

INTEGRATE

TEST

FUZZY LOGIC

STANDARD COMPONENT

SUBSYSTEM

MODULES

SCREENS

REPORTS

NEW & MODIFIED

FUNCTION POINTS

BALL PARK SIZE ESTIMATE

BASE LINE SIZE ESTIMATE

CHANGE CONTROL SIZE ESTIMATES

STANDARD COMPONENT

SUBSYSTEM

MODULES

SCREENS

REPORTS

NEW & MODIFIED

FUNCTION POINTS

HISTORY DATA BASE TO SUPPORT PLANNING

SLOC

BITS

FILES

BYTES

BATCH PGM

WORDS

ONLINE PGM

OBJECT
LOW DISPERSION AT EXPECTED VALUE INDICATES LOW UNCERTAINTY IN THE ANSWER.