

# Software Risk Assessment

## **Analyzing the Fallible Machine**

The Open PSA Initiative

## A Blending of Disciplines

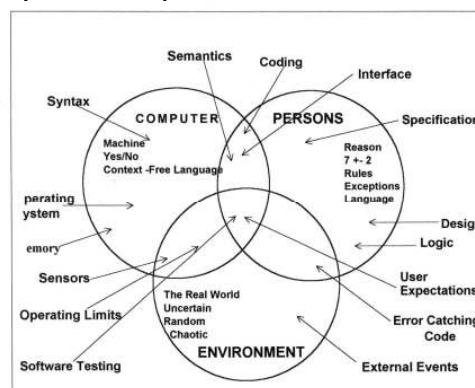
- Systems Engineering
- Software Reliability
- Software System-Safety
- Probabilistic Risk Assessment
- Software Engineering
- Software Development  
Management

## Essential Properties of Software

- Software is constructed from human thought, or reason.
- The material sciences of software are mathematics and cognitive sciences.
- Its physical existence is a set of instructions in an English-like code.
- Software does not wear out; but it does get over-maintained or over-changed.
- Each software construction is unique.
- Software is combinatorially explosive with respect to possible execution paths.
- When software fails, no deformations or breakage occur.
- Software fails with little, if any, advanced warning (we don't hear the gears grinding).
- Software errors exemplify action at a distance.

## Software System Failures

- Failures are at the intersections of system viewpoints



- Software system failures are disconnects between users, software writers, the operating environment, and the computing machine

## Systems Engineering

- Hierarchical Approach
- Software is **NOT** a Standalone Device
- Software is a Component in a System
  - Software controlled devices
  - Software supported decisions
- Concern is Primarily with Systems Effects of Software Errors
  - Therac 25
  - Precision Errors

## Software Reliability

- Software is a Product which does **NOT** perform Perfectly
- Both Parallels and Differences Between Hardware and Software Reliability
- Refers to a Collection of Concepts
  - Number of software errors
  - Software RAM
  - Software failure rate (F, R, z, H)
  - MTTF or MTBF
- Techniques which Lower Frequency and Consequences of Failures
- Measurement and Modeling of Probability of Failure

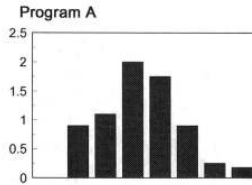
## First, Some Definitions:

- **Fault**
  - Something wrong in a program
- **Error**
  - A manifestation of a **FAULT** during software operation
- **Failure**
  - An **ERROR** which interferes with system operation
- **Bug**
  - Any of the above
- **Feature**
  - A **BUG** to which users have become accustomed

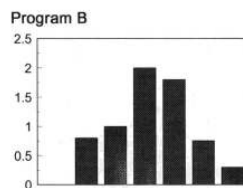
## Software Error Model Definitions

- $E_T$ 
  - Total number of errors in program
- $E_C(\tau)$ 
  - Total number of errors corrected after time  $\tau$  of debugging
- $E_R(\tau) = E_T - E_C(\tau)$ 
  - Total number of errors remaining after time  $\tau$  of debugging
- $I_T$ 
  - Total number of program instructions, or lines of code

## Error Removal Rate for Two Large Programs



Months of Testing/Debugging	0	1	2	3	4	5	6	7
Error Rate per 1000 Instructions	0.00	0.90	1.10	2.00	1.75	0.90	0.25	0.18



Months of Testing/Debugging	0	1	2	3	4	5	6
Error Rate per 1000 Instructions	0.00	0.80	1.00	2.00	1.80	0.75	0.30

## Reliability Definitions

-  $R(t) = e^{-\lambda t}$

- Constant Hazard Reliability Function

-  $MTTF = \int_0^{\infty} tf(t)dt$

- MTTF: Mean Time to Failure

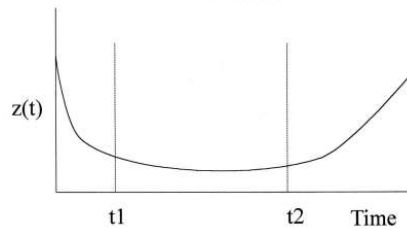
-  $MTTF = \int_0^{\infty} R(t)dt$

- Alternatively, and more simply

-  $MTTF = \int_0^{\infty} e^{-\lambda t} dt = \frac{e^{-\lambda t}}{-\lambda} \Big|_0^{\infty} = \frac{1}{\lambda}$

- For a constant hazard, we obtain the above

## The Bathtub Curve of Failure Rates

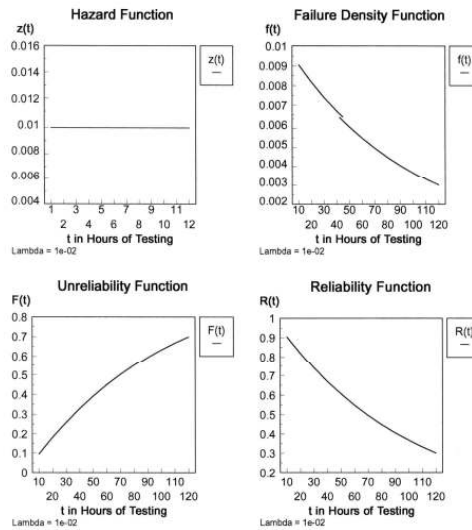


- By studying this curve for electrical or mechanical failure rates we see that:
  - Burn-in minimizes early failures
  - Scheduled maintenance minimizes late failures
  - Operational failures happen at a constant rate
- Software renames things a bit:
  - Burn-in is code release
  - Wear-out is code maintenance

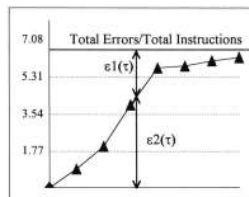
## The Constant Hazard Model

- $z(t) = \lambda$ 
  - Hazard, or failure rate, function
- $f(t) = \lambda e^{-\lambda t}$ 
  - Failure density function
- $R(t) = e^{-\lambda t}$ 
  - Reliability function
- $F(t) = 1 - R(t)$ 
  - Unreliability function

## The Constant Hazard Model



## Normalizing the Software Error Model



- Cumulative Errors per 1000 Instructions for Program A
- We Normalize the Error Rate to the Number of Instructions in a program:

$$\frac{E_R(\tau)}{I_T} = \frac{E_T}{I_T} - \frac{E_C(\tau)}{I_T}$$

- Notice that:

$$\epsilon_1(\tau) = \frac{E_R(\tau)}{I_T} \text{ and } \epsilon_2(\tau) = \frac{E_C(\tau)}{I_T}$$

## Software Failure Rate Model

- Failure Rate is Proportional to Number of Remaining Errors
  - Intuitive
  - Model is Simple
  - Musa has Data to Support

- $$z(t) = K(\varepsilon_1(\tau)) = K \left[ \frac{E_T}{I_T} - \varepsilon_2(\tau) \right] \quad !:$$

we define normalized parameters  $\alpha$  and  $\beta$ :

$$\beta = \frac{E_T}{I_T} K \quad \text{and} \quad \alpha = \frac{\varepsilon_2(\tau) I_T}{E_T}$$

so:

$$z(t) = \beta(1 - \alpha t)$$

## Example of Failure Rate Estimation

After 1 month of integration testing, a 25,000 line long program was found to have 12 errors. Since about 5 tests a day were run at 1 hour/test, the average MTTF was 8 hrs. After 2 months, 18 total errors were found with a MTTF of 12 hrs.

- Estimation of Model Parameters

$$z(\tau) = \frac{1}{\text{MTTF}} = \frac{K}{I_T} [E_T - E_C(\tau)]$$

$$.125 = \frac{K}{2.5 \times 10^4} [E_T - 12]$$

$$.083 = \frac{K}{2.5 \times 10^4} [E_T - 18]$$

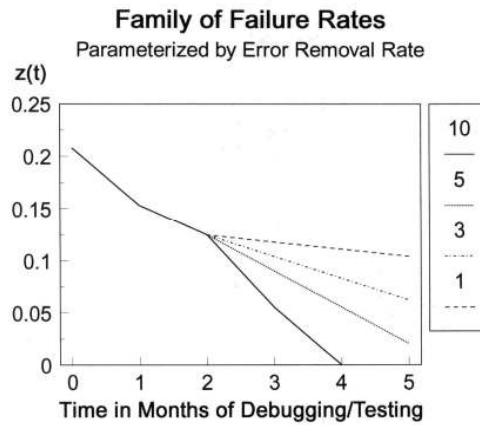
$$\frac{.125}{.083} = 1.5 \approx \frac{E_T - 12}{E_T - 18} \Rightarrow E_T = 30$$

$$.125 = \frac{K}{2.5 \times 10^4} [30 - 12] \Rightarrow K = 173.6$$

$$z(\tau) = 6.94 \times 10^{-3} [30 - E_C(\tau)]$$



## Estimation Example



Estimation will change if a different error removal rate is assumed or if it changes during testing.

## Reliability and Failure Rates

- Specification
  - Failure rates should be based on measurements during field operation of similar successful and unsuccessful software.
- Prediction
  - Early prediction requires a similar project for study and for failure rate model parameters; in short, a database.
- Measurement
  - Record running hours of program during test and simulation; count software failures; calculate failure rate during development at several points.
- Confirmation
  - Confirm, or update, model during field test and early operation.

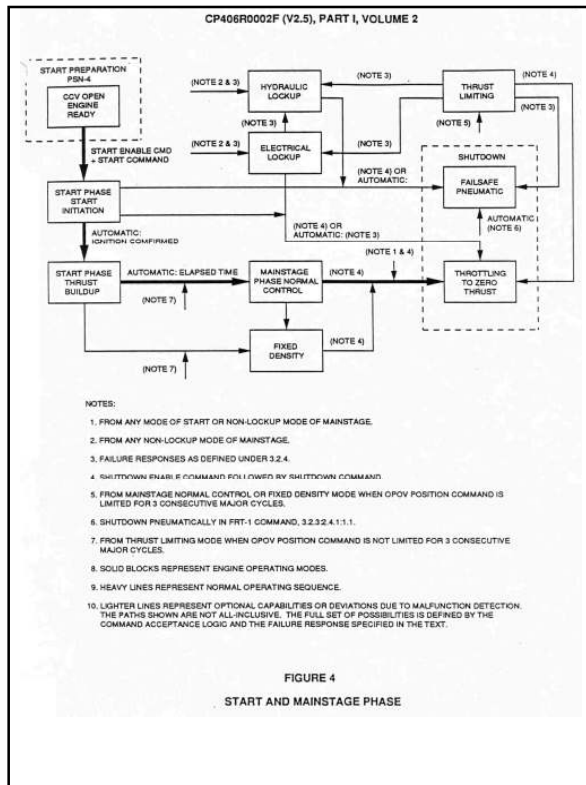
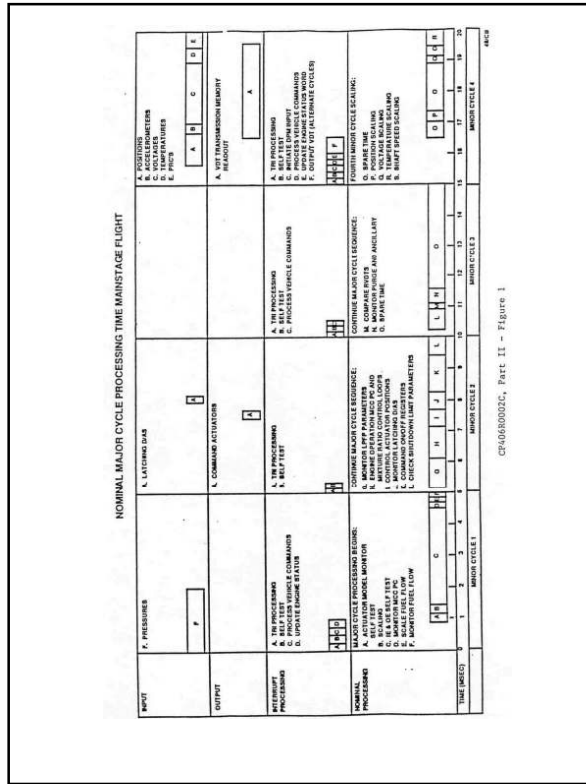
## Software System-Safety

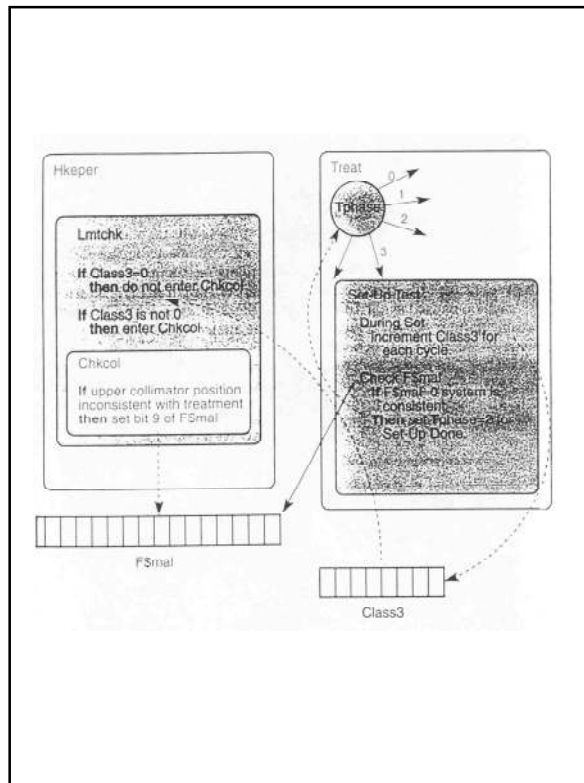
- Software Hazard Identification
- Hazard Assessment
  - System Impact
  - Consequence Criticality
  - Mitigation Possibilities
- Software Failure Mode Identification
- Design Software “devices” to Eliminate or Control Hazards

## Tools and Techniques for Software System-Safety Analysis

- Integrated into the Software Life-cycle
- Analysis is Iterative
- Code Walkthroughs
- Safety Critical Functions and Variable Analysis
- HAZOPs
- Fault Tree Analysis
- Software Root Cause Analysis







## Mathematical Proofs

- Floyd/Hoare (1975)
  - I/O Assertion Method
  - Loop Invariant Statements
  - Symbolic Execution
  - Hand/Automated Proofs
  - Assertions and Invariants Extremely Difficult to Formulate
- D. Parnas (1991)
  - Rewrite Requirements into A-7 Event and Condition Tables
  - Handproof using Functional Abstraction called Program Function Tables
  - Only Proves that Program Fulfills Requirements
  - At Ontario-Hydro, 7,000 LOC took 30 man/years
- Proof Procedure is as much in question as the Code

# Safety Critical Variable Analysis

**Variable Name:** Oxidzer\_Coeff0

**Definition:** Oxidizer coefficient; loaded with adaptation data

**Units:** psia/sec/lb      **Minimum:** na      **Maximum:** na

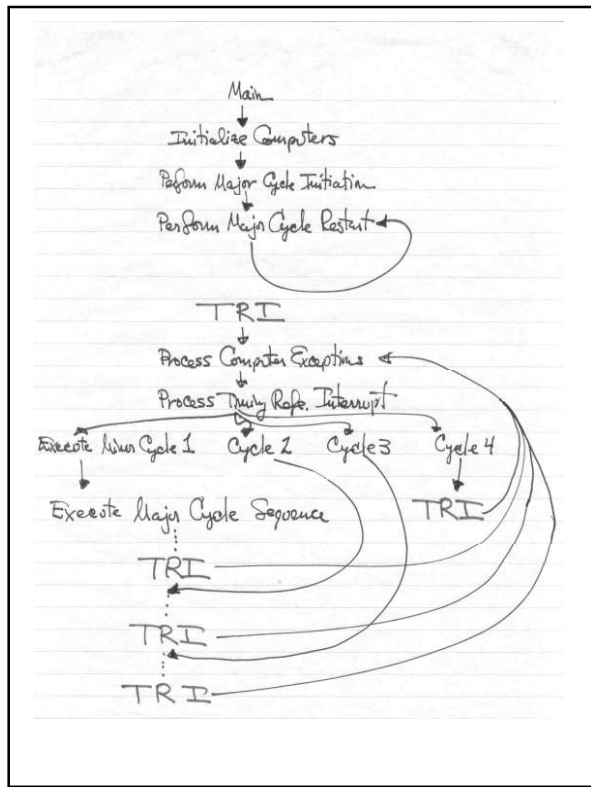
**Data Type:** Constant      **Extent:** Global      **Value:** 2.92104

**Function Name:** Calculate\_Mixture\_Ratio (285)

Variable Affected	Line	How Affected	Function Performed	Candidate for Analysis
Oxidizer_Flow_Rate_Coeff	17	Assignment	f(Oxidizer_Coeff0)	no
Lox_Mass_Flow	18	Assignment	f(Oxidizer_Flow_Rate_Coeff)	no
Oxidizer_Flow_Rate	19	Assignment	f(Lox_Mass_Flow)	no
Mixture_Ratio	23	Assignment	f(Oxidizer_Flow_Rate)	yes
Mixture_Ratio	22	Boolean	f(Lox_Mass_Flow)'	yes

<sup>1</sup> If Lox\_Mass\_Flow is >>> then Mixture\_Ratio is set to 15.999511719

Note: Mixture\_Ratio is the only variable which needs trace continued for assessing impacts of Oxidizer\_Coeff0



Form PEG-021 P.T.G. Inc. SFA

Job No: Job Title: SFA

Task/Subject:

Prepared By: Date: Checked By: Date:

### TRACE of Oxidizer Coeff

#### SEGMENTS

① Calculate Mixture Ratio 285

Line	Variable	Equation
17	Oxidizer_Flow_Rate_Coeff	$= f(\text{Oxidizer\_Coeff})$
18	LOX_Mass_Flow	$= f(\text{Oxidizer\_Flow\_Rate\_Coeff})$
19	Oxidizer_Flow_Rate	$= f(\text{LOX\_Mass\_Flow})$
23	Mixture_Ratio	$= f(\text{Oxidizer\_Flow\_Rate})$

Note: that Oxidizer\_Flow\_Rate is never used or set again in OP. Why?? (ans: it is in the VDI)

Limit Checks: if LOX\_Mass\_Flow is very large the calculation for Mixture\_Ratio might overflow. If so Mixture\_Ratio is set to a coded value of 15.99971713

Since Mixture\_Ratio is GLOBAL, we must continue the TRACE

### TRACE of Mixture Ratio

#### SEGMENTS

① Perform MR Control loop 296

Line	Variable	Equation
12	MR_Integrator_Input	$= f(\text{Mixture\_Ratio})$
13	MR_Proportional_Output	$= f(\text{MR\_Integrator\_Input})$
22, 27	MR_Integrator_Input_Past	$= f(\text{MR\_Integrator\_Input})$
23, 26	MR_Integrator_Output	$= f(\text{MR\_Integrator\_Input\_Past}, \text{MR\_Integrator\_Input})$

Note: (a) MR\_Integrator\_Output depends on this and previous Major Cycle's value of MR\_Integrator\_Input; GLOBAL, but only used locally.

(b) Both MR\_Proportional\_Output and MR\_Integrator\_Output are GLOBAL, the TRACE continues

Form PEG-021 P.T.G. Inc. SFA

Job No: Job Title: SFA

Task/Subject:

Prepared By: Date: Checked By: Date:

### TRACE of MR Proportional Output MR\_Integrator\_Output

#### SEGMENTS

- ① Generate FPOV Commands 303
- ② Initialize Control loop 290
- ③ Initialize Start Cycle 259
- ④ Perform Subsystem The Forces 269
- ⑤ Perform Start/Run transition 261

① Generate FPOV Commands

Line	Variable	Equation
12	FPOV_Goal	$= f(\text{MR\_Proportional\_Output}, \text{MR\_Integrator\_Output})$

Limit Check: if FPOV\_Goal > FPOV\_Cycle\_Limit, then FPOV\_Goal is set to that limit.

Note: FPOV\_Goal is used in Perform MR Control Loop, but not in a way that affects other variables like those already being considered.

② Initialize Control loop

Line	Variable	Equation
12	MR_Integrator_Output	$= f(\text{MR\_Proportional\_Output})$

Note: nothing new will be added to the TRACE table.

③ Initialize Start Cycle

Line	Variable	Equation
12	MR_Integrator_Output	$= f(\text{MR\_Proportional\_Output})$
13	MR_Proportional_Output	$= f(\text{MR\_Integrator_Output})$

Note: these variables are initialized only

Form P1-G-721 PLG, Inc.           
 Job No:          Job Title: SFA  
 Task/Subject:           
 Prepared By:          Date:          Checked By:          Date:         

① Perform Shutdown Time Function  
 line 3, 10 Same as Initialize Start Byte

② Perform Start to Mainstage Transition  
 line 18  $MR\_Interpctn\_Output = f(MR\_Propctnl\_Output)$   
 Note: Again nothing new added to path

---

Summary

SEQUENTS

- ① Calculate Mixture Ratio
- ② Perform MR Control Logic
- ③ Generate FPOV Commands
- ④ Initialize Control Loop
- ⑤ Perform Shutdown Time Function
- ⑥ Initialize Start Byte
- ⑦ Perform Start to Mainstage Transition

Variable Dependency

Form P1-G-721 PLG, Inc.           
 Job No:          Job Title: SFA  
 Task/Subject:           
 Prepared By:          Date:          Checked By:          Date:         

But that's not the end of it. We must make sure that none of the variables in the trace have external names equated to them.

FPOV\_Goal has aliases!

- ① FPOV\_Goal EQUATES TO FPOV\_Goal\_Ch\_A
- ② FPOV\_Goal is a child of Actuator\_Goal
- ③ FPOV\_Goal\_Ch\_A is a child of Actuator\_Goal\_Ch\_A
- ④ Actuator\_Goal\_Ch\_A is a child of Actuator\_Goal

SEQUENT  
 REQUEST to Ramp 741  
 line Actuator\_Goal is set to Act\_Goal  
 Note: no interactions or dependencies

SEQUENT  
 TRIP of Actuator\_Goal\_Ch\_A  
 Control Actuator Positions 465  
 line 14 If Actuator\_Goal\_Ch\_A  $\neq$  Actuator\_Goal\_Ch\_A  
 ① Compute new value  
 ② Set new value flag

52 If new value  
 ① Send output: ~~ISSUE~~ OF Commands  
 → \*CHANGE in Value\* ←



# Software Hazard Analysis Report

Identification No. \_\_\_\_\_

Title

Software Affected

Mission Phase/Major Mode

Hazard Description/Discussion

Recommended Hazard Controls

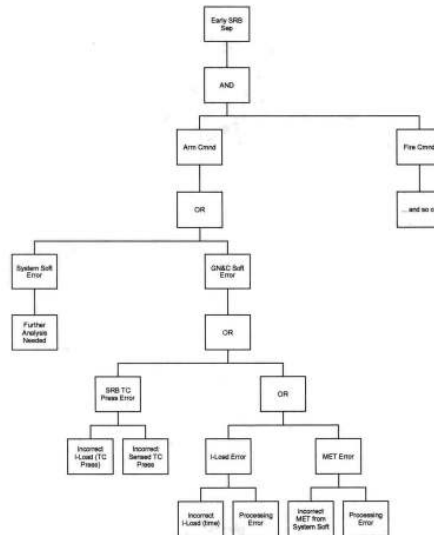
Action/Status/Disposition

Hazard Tracking Documents

As Of Date \_\_\_\_\_ Status \_\_\_\_\_ Originator \_\_\_\_\_ Date \_\_\_\_\_

Closure Sign-off \_\_\_\_\_ Date \_\_\_\_\_

## Software Fault Tree



## Software Error Root Cause Analysis

- For Each Project the Following Data Should be Kept
  - Life-cycle used
  - Start and end date for each phase
  - Effort spent in each phase
  - Development Environment
  - Programming experience
  - SEI Index
  - Number of lines of code
  - Source language
  - Target hardware
- For Each Failure, a Summary and Data Record Should be Created

## Software Error Root Cause Analysis Report

SERCA Number:

Date/Time:

Severity of Failure:

Failure Description:

Activity Being Performed:

Method of Detection:

Type of Failure:

Unit Complexity:

Unit Size:

Fill in AT LEAST ONE of the following:

CPU hours since last failure:

Runs since last failure:

Calendar hours since last failure:

Labor hours since last failure:

Hours/Interval and number of errors in this interval:

Type of software fix:

Fill in AT LEAST ONE of the following:

CPU hours required to fix:

Runs to fix:

Calendar hours to fix:

## Some Observations

- |   |  |
|---|--|
| <ul style="list-style-type: none"><li>▪ Software Reliability</li><li>– Quantifies Failure Rate</li><li>– Does <b>NOT</b> look at code</li><li>– Relies on Measurement</li><li>– Applies to the MACRO level</li><li>– Results are "Hard"</li></ul> | <ul style="list-style-type: none"><li>▪ Software System-Safety</li><li>– Does <b>NOT</b> Quantify</li><li>– Analyzes the code</li><li>– Relies on Judgment</li><li>– Applies Hierarchically</li><li>– Results are "Soft"</li></ul> |
|---|--|

## The Trouble with Testing

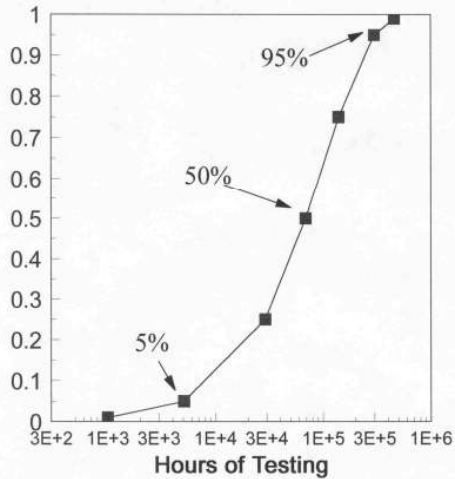
- Hecht's Law (1992)
  - Infrequently Executed Code has a High Failure Rate
    - Redundancy Management
    - Exception Handling
    - Initialization
    - Calibration
  - Off-nominal Conditions are a Prominent Cause of Failure in Well-tested Systems
  - Test Profile must be Rich in Off-nominal Conditions
  - Software Telemetry Needed
- The Butler/Finelli Observations (1991)
  - Reliability Quantification to Low Levels is Statistically Infeasible
  - Separately Programmed Versions do **NOT** Fail Independently

## A Testing Example

### Confidence in Software Failure Rate

$$z(t) = 1e-05/\text{hr.}$$

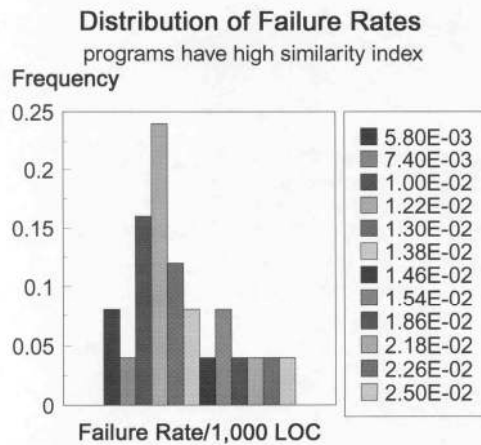
Level of Confidence



## Software Risk Assessment Methodology

- Elements of PSA
  - Uncertainties
  - Bayes' Theorem
  - Scenarios
    - Master Logic Diagrams (MLD)
    - Event Sequence Diagrams (ESD)
    - Initiating Event Fault Trees
  - Propagation of Uncertainties
- Synthesis into a Programmatic Approach
  - Hierarchical Analysis
  - Quantitative Methods from Software Reliability
  - Qualitative Methods from Software System-Safety

## Uncertainties about Failure Rates



## Criteria for Calculation of an Application Similarity Index

- Design
- Operational Profile
- Type of Project
- Size of Code
- Complexity Metrics
- Development Environment
- Language Used
- Total Test Time

Similarity index is calculated by solving for the principal eigenvector of an application comparison matrix.

## Complexity Metrics

- Measures
  - Size of Code
  - Number of Branches
  - Program Structure
  - Flow of Control
- Used on the Subprogram Level
- Commonly Used Metrics
  - SLOC
  - McCabe's Cyclomatic Complexity
  - NPATH
  - Halstead Software Science
- Highly Correlated with Each other
- Moderate Correlation between Complexity Thresholds and Failure Rates

```

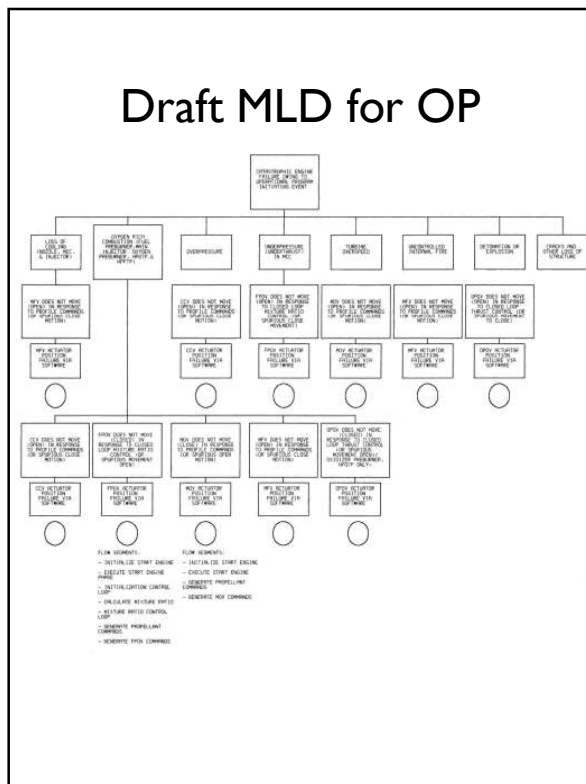
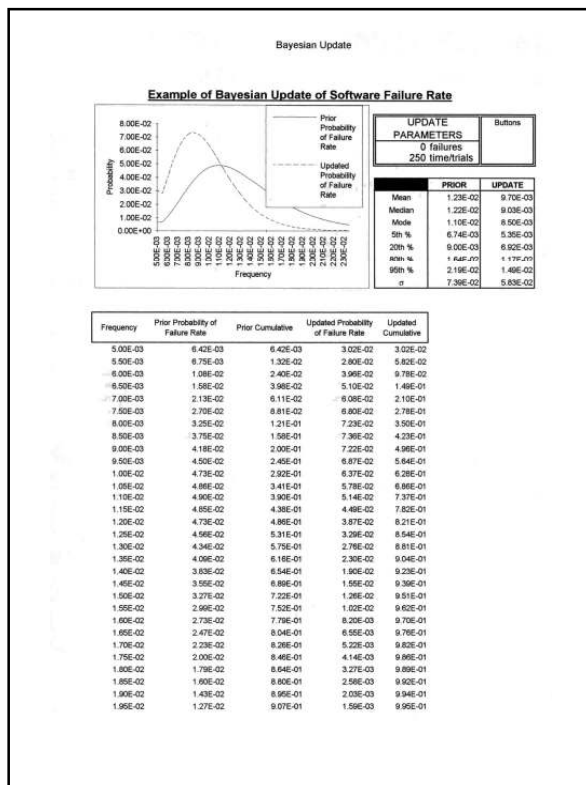
Number of Ada Files measured : 959      F05BZ1
-----
n1 (total) : 12814      n1 (average) : 13.4
n2 (total) : 24127      n2 (average) : 25.2
n (total) : 36941      n (average) : 38.5
N1 (total) : 81571      N1 (average) : 85.1
N2 (total) : 60125      N2 (average) : 62.7
N (total) : 141696      N (average) : 147.8
N^ (total) : 189260.8   N^ (average) : 197.4
V (total) : 943270.9    V (average) : 983.6
ALOC (total) : 19692     ALOC (average) : 20.4
Ss (total) : 27639      Ss (average) : 28.8
TLOC (total) : 113670    TLOC (average) : 118.5
C (total) : 4188.3      C (average) : 4.4
VARS (total) : 17930     VARS (average) : 18.7
v(G) (total) : 4401      v(G) (average) : 4.6
Lambda (total) : 4394.5  Lambda (average) : 4.6
T^ (total) : 558.5      T^ (average) : 0.6
-----

```

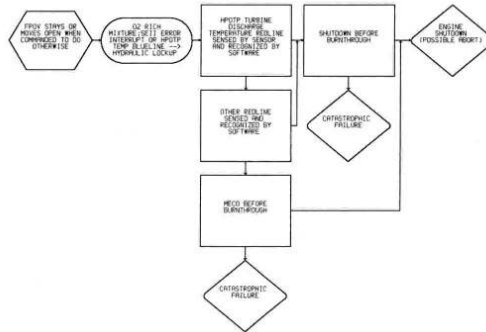
```

Ada Software Metrics Descriptions:
-----
n1 - Number of Unique Operators
n2 - Number of Unique Operands
R - Vocabulary (Number of Unique Operators + Number of Unique Operands)
N1 - Total Number of Operator Occurrences
N2 - Total Number of Operand Occurrences
N - Size (Token count = Total Number of Operator + Operand Occurrences)
N^ - Estimated Program Length
    (Length Equation = (n1 x log2(n1) + n2 x log2(n2)))
V - Volume (in bits = N x log2(n), A program needs approximately
    log2(n) bits to represent the Vocabulary n)
ALOC - Ada Lines of Code (counts 1 Ada line of code per semi-colon)
Ss - Lines of Code (counts all non-blank and non-comment lines)
TLOC - Total Lines of Code (all lines of code including blank and comment li
    nes)
C - Language Constant (N / Ss (for Fortran C is approximately 7.0))
VARS - Number of Unique Variables (n2 - unique constants - labels)
v(G) - Cyclomatic Complexity (McCabe's metric - number of distinct basic
    paths through a program)
Lambda - Language Level Constant
T^ - Estimated Programming Time (in man-hours) = Effort/Beta (where
    Beta = Stroud's Number = number of elementary mental
    discriminations/second)
-----

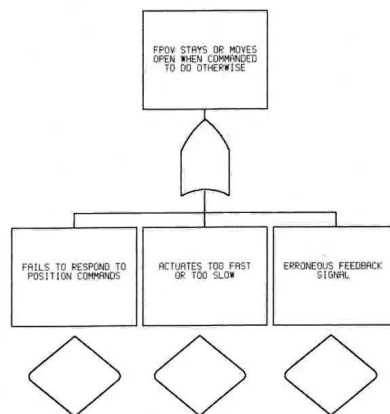
```



## Draft Engine Level ESD for OP Initiated Scenario Main Stage Closed Loop Control



## Draft Example Fault Tree for Initiating Event





## Propagation of Uncertainties

- Event Sequence Diagrams are Used to Construct Event Trees
- Pivotal Events become Top Events
- Software Pivotal Events are Intermixed with Phenomenal and Failure Pivotal Events
- Each Path through an Event Tree Depicts a Possible Scenario
- Probability Distributions for each Scenario are created by Monte Carlo Simulation

## Steps in an SRA

- Review Software
  - Purpose
  - Functions
  - Construction
  - Operation
  - Software Engineering
- Review Controls
  - Reliability Models
  - Failure Data
  - Problem Reports
  - Testing Program
- Develop Scenarios
  - Add to Existing or Create New Scenarios
  - Employ
    - MLDs
    - ESDs
    - SFTs
    - Event Trees

## Steps in an SRA (Continued)

- **Collect Data**
  - Four types of data
    - Operational failure data and time
    - Test data and time
    - Failure Rate from similar applications
    - Judgment from expertise and experience
  - Develop prior distributions
  - Modify with test data
  - Develop likelihood functions
  - Create application specific data
  - Probability (Critical Failure | Failure)
- **Allocate failure rate over the software functions**
  - By fraction of time during nominal processing
  - By Proportion of Similar Failures
  - By Complexity of Functions in Scenario
  - By Expert Elicitation
- **Quantify Scenarios**