



# Touchstone

**Assurance for NextGen Software-Intensive Systems**

**Chuck Howell, [howell@mitre.org](mailto:howell@mitre.org)**

**Gary Vecellio, [vecellio@mitre.org](mailto:vecellio@mitre.org)**

**November 2008**

# What we do NOT want



# The Problem: Current Challenges

**“Certification of software-intensive systems today is largely based on adherence to specified processes, rather than on a systematic evaluation of the safety and quality of the systems themselves. This process-based approach to certification is likely a primary reason for excessive costs and time. An effective program must have as one of its primary goals the replacement of the process-based approach with an approach (or approaches) based on a systematic evaluation of the systems themselves.”**

**“Visions of Automation and Realities of Certification”, Hayhurst and Holloway, AIAA 5th Aviation, Technology, Integration, and Operations Conference 26-29 September 2005**

# Example: DO-178B

## 9.0 CERTIFICATION LIAISON PROCESS

### 9.2 Compliance Substantiation

The applicant provides evidence that the software life cycle processes satisfy the software plans. Certification authority reviews may take place at the applicant's facilities or the applicant's suppliers' facilities. This may involve discussions with the applicant or its suppliers. The applicant arranges these reviews of the activities of the software life cycle processes and makes software life cycle data available as needed.

#### DO-178B definitions:

“One distinction between reviews and analyses is that analyses provide repeatable evidence of correctness and reviews provide a qualitative assessment of correctness. A review may consist of an inspection of an output of a process guided by a checklist or similar aid.”

# Getting philosophical



## HAZARDS

THERE IS AN ISLAND OF OPPORTUNITY IN THE MIDDLE OF EVERY DIFFICULTY.  
MISS THAT, THOUGH, AND YOU'RE PRETTY MUCH DOOMED.

[www.despair.com](http://www.despair.com)

# Proof by Repeated Assertion

**“Few, if any, existing certification regimes encompass the combination of characteristics recommended in this report – namely, explicit dependability claims, evidence for these claims, and a rigorous argument that demonstrates that the evidence is sufficient to establish the validity of the claims”.**

***Software for Dependable Systems: Sufficient Evidence?* U.S. National Research Council, 2007**

# Current Support for Dependability Cases is not Adequate

- **Notation and methods help, but...**
- **Recent illustrations using GSN**
  - Manual analysis exposed flaws in arguments
  - Anecdotal response: “they are in GSN, they must be correct”
- **Examples from other domains strongly suggest that “expert judgment” model is much worse than we like to believe**

# Analysts: Lessons from Cognitive Science

*Based on material by Dr. Paul Lehner, MITRE*

- **Over 130 studies, covering 40+ years of research in many domains, suggest that:**
  - **People are good at picking out the right predictor variables**
  - **People are bad at integrating information from diverse sources.**
  
- **Experts and novices alike are subject to natural cognitive limits and judgment biases**
  - **When evaluating evidence**
  - **When reasoning about cause and effect**
  - **When estimating probabilities**

# Project Goals (50,000 ft)

- **Evaluate tools and techniques for the development, approval, and maintenance of *dependability cases*\* as engineering artifacts**
  - Use ASCE with schema extensions and plugins
  - Apply techniques to enhance review of dependability cases as engineering artifacts
- **Incrementally enhance scrutiny, rigor, workflow, collaboration, efficiency, effectiveness**

\*Adopting term from *Software for Dependable Systems: Sufficient Evidence?*  
U.S. National Research Council, 2007

# Some Techniques to be Adapted

- **Software inspections**
  - Perspective based reading, phased inspections
- **Tim Kelley’s “reviewing assurance arguments”**
  - Argument comprehension, well formed structurally, expressive sufficiency, argument criticism and defeat
- **W. Greenwell’s “taxonomy of argument fallacies”**
  - Circular reasoning, mathematical fallacies, unsupported assertions, anecdotal arguments, omission of key evidence, linguistic fallacies,... 8 classes, 33 total
- **Legal reasoning analysis of “defeasible arguments” and attacks on arguments**
  - Undercutting, rebutting, disputing inductive premise,...

# Dependability Cases as Engineering Artifacts

## Why Scrutiny (disciplined collaborative inspection)?

“Engineers today, like Galileo three and a half centuries ago, are not superhuman. They make mistakes in their assumptions, in their calculations, in their conclusions. That they make mistakes is forgivable; that they catch them is imperative. Thus it is the essence of modern engineering not only to be able to check one’s own work, but also to have one’s work checked and to be able to check the work of others.”

H. Petroski in *To Engineer is Human: The Role of Failure in Successful Design*

## Why Rigor (automated analysis)?

“The first principle is that you must not fool yourself, and you are the easiest person to fool.”

Richard P. Feynman (credit for this quotation due to Alloy tutorial)

# Project Plan

- **Start with existing notations and tools**
  - Extend cia plug-ins and schema extensions
- **Start with sample FAA safety review as case study (prefer DO-278 to DO-178B)**
  - Map interim safety case to dependability case, validate mapping
  - Demonstrate tools and techniques for rigor and scrutiny in reviews of dependability cases
  - Did it help? Was it worth it? Reason to believe it scales?
- **Follow on with incremental extension for rigor, more support for scrutiny**

# Hurdles to adoption

- **“It is quite possible to follow a faulty analytical process and write a clear and persuasive argument in support of an erroneous judgment.”**  
R. Heuer, *The Psychology of Intelligence Analysis*
- **“Nothing will ever be attempted if all possible objections must first be overcome.”**  
Samuel Johnson
- **“It is difficult to get a man to understand something when his salary depends upon his not understanding it.”**  
Upton Sinclair

# Limits of Tools and Techniques

They're teaching a new way of plowing over at the Grange tonight - you going?

Naw - I already don't plow as good as I know how...



“Knowing is not enough, we must apply. Willing is not enough, we must do.”  
Goethe

# “But enough about our project...”

- **We want to hear from you – what do *you* think of our project 😊?**
  - **How do you approach review of dependability cases?**
  - **How do you use ASCE in reviews?**
    - **What extensions do you use, would you find useful?**
  - **Suggestions for tools, techniques, notations to look at?**
  - **Other relevant work we should become familiar with?**
  - **Please contact us via email with any suggestions or questions**

# THANK YOU!



## QUALITY

THE RACE FOR QUALITY HAS NO FINISH LINE-  
SO TECHNICALLY IT'S MORE LIKE A DEATH MARCH.

[www.despair.com](http://www.despair.com)

# Backup Material



# **“The Trouble with Dependability Cases Today” ...**

- **Much room for improvement**
  - **Post-mortems of spectacular mishaps**
  - **Cost and burden of high assurance systems**
- **“It’s not how hard you worked, it’s what you accomplished”**
- **There are problems in every aspects of dependability cases**
- **The volume of material combined with little structuring support and ad hoc “rules of evidence” are the root of many problems**

# Building the Dependability Case

- **Most guidance is strong on excruciating detail for format, weak on gathering, merging, and reviewing technical evidence**
- **Guidance often uses the “cast a wide net” tactic**
  - Assurance costs time and money
  - “Squandered diagnostic resources”
  - Some work on a “portfolio management” approach to assurance

# Reviewing the Dependability Case

- **Stacks of free-format text makes the review process tedious**
  - Hard to see linkages or patterns
  - Hides key results in sheer volume
- **Weak guidance on review of arguments and evidence often results in ad hoc criteria (be very nice to your reviewer!)**
- **Rarely is there explicit guidance for weighing conflicting or inconsistent evidence**

# Maintaining the Dependability Case

- **The one thing more brittle than software: the associated dependability case**
- **Volume and relatively rare use of tools make it difficult to understand impact of a change on assurance structure**
  - **Have the claims changed?**
  - **Are arguments invalidated or new ones needed?**
  - **Is evidence still relevant, new evidence needed?**
  - **“Weak link effect” of discrete systems compounds the problem**
- **Revalidation costs are a major burden for critical systems**
- **“Breakage” when bow-waving requirements another example**