Reliability of 1oo2 Software-based Systems in which one Channel is "Possibly Perfect"

Bev Littlewood

Centre for Software Reliability, City University, London

John Rushby

Computer Science Laboratory, SRI, Menlo Park, California

The set-up

- 1-out-of-2 design-diverse, 2-channel software based system
- We are interested in *probability of failure on demand* (*pfd*)
  - E.g. reactor protection system
  - But much of what we say here may also apply in wider contexts, e.g. continuously operating fault tolerant systems
- We know such fault tolerant approaches can be effective ways to *achieve* reliability
  - E.g. reliability in eventual operational use of the Airbus A320 and later fault tolerant flight control systems?
- BUT…..

….There's a problem:

Although such an approach may work "on average" (in some sense), it's hard to know whether it has worked in a particular instance - and *how reliable* the resulting system will be in operation

- Cannot assume independence of version/channel failures
  - In fact they will *not* fail independently
- $Pfd_{\mathrm{sys}} > pfd_A . pfd_B$
  - Experiments tell us this
  - So does theory
- Need to know "how dependent" the failure processes of the different channels are
- Measuring this is as hard as measuring Pfd_{sys} by treating it as a black box
- So…an impasse?

A possible way out

Consider a 1oo2 system in which channel A is "highly functional", and therefore complex, *but channel* B *is simpler and thus possibly "perfect"*

- Perfect means it will never experience a failure
- Possibly perfect means there is some uncertainty about its perfection
  - In particular there is a probability of imperfection
- For A our uncertainty concerns whether it will fail on a randomly selected demand: probability pfd_A
- for B our uncertainty concerns whether it is not perfect: probability pnp_A

Aleatory and Epistemic Uncertainty

- Aleatory uncertainty is "uncertainty in the world", or irreducible uncertainty
  - Uncertainty about *A* failing, about *B* not being perfect - both involve aleatoric uncertainty
- Epistemic uncertainty is "uncertainty about the world", or reducible uncertainty
  - Sometimes called "model uncertainty"
  - E.g. uncertainty about the size of pfd_A and of pnp_B
- We now analyse our system in two stages: aleatoric, then epistemic

But now suppose for the moment we *know* $pfd_A = p_A$ and $pnp_B = p_B$…

Aleatoric uncertainty for 1oo2 system

$$P(\text{system fails on randomly selected demand} \mid pfd_A = p_A, pnp_B = p_B)$$
$$= P(\text{system fails} \mid A \text{ fails, } B \text{ not perfect}, pfd_A = p_A, pnp_B = p_B)$$
$$\times P(A \text{ fails, } B \text{ not perfect} \mid pfd_A = p_A, pnp_B = p_B)$$
$$+ P(\text{system fails} \mid A \text{ succeeds, } B \text{ not perfect}, pfd_A = p_A, pnp_B = p_B)$$
$$\times P(A \text{ succeeds, } B \text{ not perfect} \mid pfd_A = p_A, pnp_B = p_B)$$
$$+ P(\text{system fails} \mid A \text{ fails, } B \text{ perfect}, pfd_A = p_A, pnp_B = p_B)$$
$$\times P(A \text{ fails, } B \text{ perfect} \mid pfd_A = p_A, pnp_B = p_B)$$
$$+ P(\text{system fails} \mid A \text{ succeeds, } B \text{ perfect}, pfd_A = p_A, pnp_B = p_B)$$
$$\times P(A \text{ succeeds, } B \text{ perfect} \mid pfd_A = p_A, pnp_B = p_B)$$

Assume, conservatively, that if *B* is imperfect it fails whenever *A* does

$$P(\text{system fails on randomly selected demand} \mid pfd_A = p_A, pnp_B = p_B)$$
$$= P(A \text{ fails, } B \text{ not perfect} \mid pfd_A = p_A, pnp_B = p_B)$$

Aleatory uncertainty (contd)

$$P(A\,\text{fails}, B\,\text{imperfect} \mid pfd_A = p_A, pnp_B = p_B)$$

$$= P(A\,\text{fails} \mid B\,\text{imperfect}, pfd_A = p_A, pnp_B = p_B) \times P(B\,\text{imperfect} \mid pfd_A = p_A, pnp_B = p_B)$$

(Im)perfection of B tells us nothing about the failure of A on this demand; hence,

$$= P(A\,\text{fails} \mid pfd_A = p_A, pnp_B = p_B) \times P(B\,\text{imperfect} \mid pfd_A = p_A, pnp_B = p_B)$$

$$= p_A \times p_B$$

Compare with two (un)reliable channels, where failure of B on this demand does increase likelihood A will fail on same demand

$$P(A\,\text{fails} \mid B\,\text{fails}, pfd_A = p_A, pfd_B = p_B) \geq P(A\,\text{fails} \mid pfd_A = p_A, pfd_B = p_B)$$

Epistemic uncertainty for 1oo2 system

- We have shown that the events "*A* fails" and "*B* is imperfect" are conditionally independent at the aleatoric level
  - Probability of system failure is (conditionally) $p_A \times p_B$
- Remaining uncertainty centres *only* on p_A and p_B
- We represent this *epistemic uncertainty by*

$$F(p_A, p_B) = P(pfd_A < p_A, pnp_B < p_B)$$

  - E.g. could think of this as his Bayesian posterior distribution when an assessor has collected evidence from testing,verification, other kinds of analysis, etc, etc
- The unconditional (subjective) probability of system failure is

$$\int_{\substack{0 \le p_A \le 1 \\ 0 \le p_B \le 1}} p_A p_B \, dF(p_A, p_B)$$

Epistemic uncertainty (contd)

- The *only* source of dependence in the model comes in via *F*
- If this were to factorise, i.e assessor's beliefs about the parameters were independent,

$$\text{P(system fails on randomly selected demand)} = P(A \text{ fails}, B \text{ not perfect})$$

$$= \int_{\substack{0 \le p_A \le 1 \\ 0 \le p_B \le 1}} p_A p_B \, dF(p_A, p_B)$$

$$= \int_{0 \le p_A \le 1} p_A \, dF(p_A) \times \int_{0 \le p_B \le 1} p_B \, dF(p_B)$$

And the assessor's task is reduced to estimating just the two posterior (marginal) means

- But this will never be true!

Reliability estimation of 1oo2 system

- Most assessors would find it hard to tell us what their *F* is
- So what can be done?
- Well….where does the "dependency of beliefs" about the parameters come from?
- A source of dependency is the possibility of common faults at a high level, e.g. misunderstanding of system requirements
- One way forward is to place probability mass, say *C*, at the point (1,1) in the (p_A,p_B)-plane to represent the assessor's (subjective) probability that there *are* such faults
- The effect of this is *conservative*: if there *are* such faults he believes *A* fails with certainty, and *B* is not perfect with certainty
  - there is a chance *C* that $p_A.p_B=1$, i.e. that the system is certain to fail

Reliability estimation (contd)

$$P(\text{system fails on randomly selected demand}) = \int_{\substack{0 \le p_A \le 1 \\ 0 \le p_B \le 1}} p_A p_B dF(p_A, p_B)$$

$$= C \times \int p_A p_B dF(p_A, p_B \mid p_A = p_B = 1) + (1 - C) \times \int p_A p_B dF(p_A, p_B \mid p_A, p_B \neq 1)$$

$$= C + (1 - C) \times \int p_A p_B dF(p_A, p_B \mid p_A, p_B \neq 1)$$

But the last integrand factorises, so

$$P(\text{system fails on randomly selected demand})$$

$$= C + (1 - C) \times \int p_A dF(p_A \mid p_A \neq 1) \times \int p_B dF(p_B \mid p_B \neq 1)$$

$$= C + (1 - C) \times P_A^* \times P_B^*$$

Discussion

Of course this is not a silver bullet. But…

- The handling of aleatory uncertainty is greatly simplified compared with the case of two *certainly fallible* channels
- The architecture *is* a special one, but it is very plausible for certain applications
  - E.g. as a means of *achieving* reliability for, say, a protection system; or for functional channel plus monitor; or highly functional channel plus get-you-home channel
- The conservative bottom-line result involves only *three parameters* and it may be possible to estimate these for real systems