CITY UNIVERSITY LONDON

Safety and Assurance Cases: current practice and the challenge of complex open systems

DEOS, Tokyo Dec 2010

Robin E Bloomfield
Adelard LLP and CSR City University London

reb@adelard.com
reb@csr.city.sc.uk
College Building, City University, London EC1V 0HB
Tel: +44 20 7490 9450 (sec Adelard)
Tel: +44 20 7040 8420 (sec CSR)

CSR Building confidence in a computerised world
www.csr.city.ac.uk

Adelard

Overview

- Introduction
- Safety and assurance cases
- Outline of research landscape
- The challenge of complex systems
- Conclusions and discussions

CSR Building confidence in a computerised world
www.csr.city.ac.uk

Adelard

Wednesday, 15 December 2010

Adelard

- Safety and assurance cases and safety management systems
- Independent safety assessment
- Software assurance, including formal methods and static analysis
- Development, interpretation and application of standards and guidelines
- applied research in safety, security, critical infrastructure interdependencies
- policy to technology
- ASCE – the Assurance and Safety Case Environment
- clients in nuclear, defence, financial, transport sectors

Centre for Software Reliability

- Evaluation of socio-technical systems
  - Technical, interdisciplinary
- Research
  - with international community and users
- Education
  - placements, internships, scholarships, courses, MSc and CPD
- Innovation
  - director, Dr Peter Popov
  - DivSQL, PIA-FARA

In the beginning…

- “The World, according to the best geographers, is divided into Europe, Asia, Africa, America, and Romney Marsh",

wrote the Reverend Richard Harris Barham, writing as Thomas Ingoldsby, in the 1840s.

Some Definitions

"A documented body of evidence that provides a convincing and valid argument that ... adequately safe for ... env...

A structured **argument**, supported by a body of **evidence**, that provides a compelling, comprehensible and valid case that a **system is safe** for a given application in a given ...

A security assurance case ... and a correspond... system s... properti...

A f... and ... that ... railwa... the sa... ...assurance ...r change to the ...fety requirements and that ...uirements are adequate.

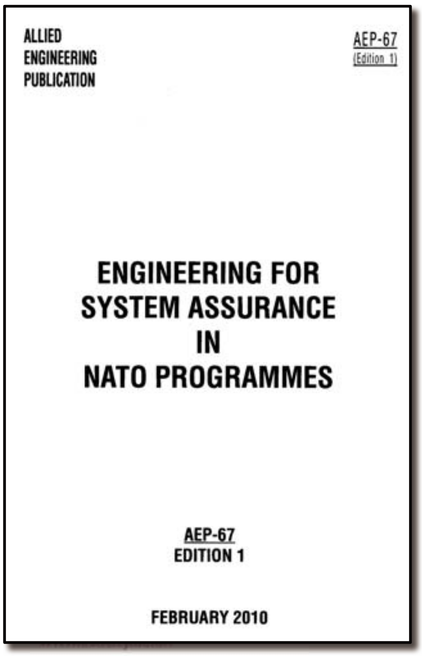
Yellow Book issue 4

An assurance case is reasoned, auditable artefact created to support the contention its claim or claims are satisfied. It contains the following and their relationships:
One or more claims about properties.
Arguments that logically link the evidence and any assumptions to the claim(s).
A body of evidence and possibly assumptions supporting these arguments for the claim(s).

ISO 15026

Supply chains

- evaluation
- communication

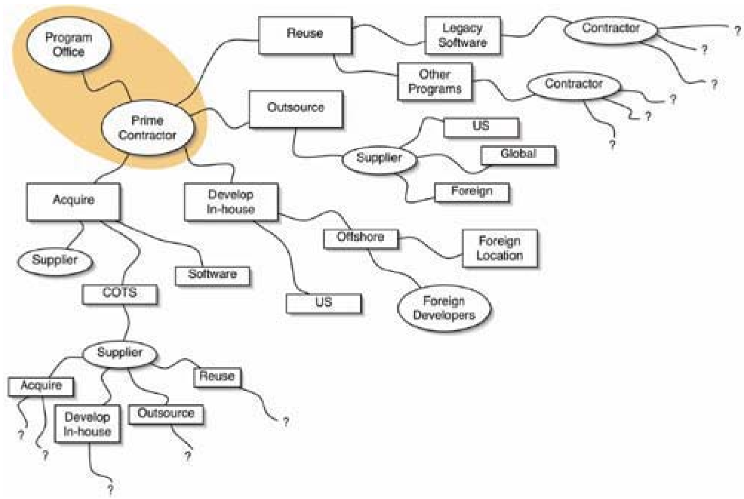

ALLIED ENGINEERING PUBLICATION

AEP-67 (Edition 1)

ENGINEERING FOR SYSTEM ASSURANCE IN NATO PROGRAMMES

AEP-67
EDITION 1

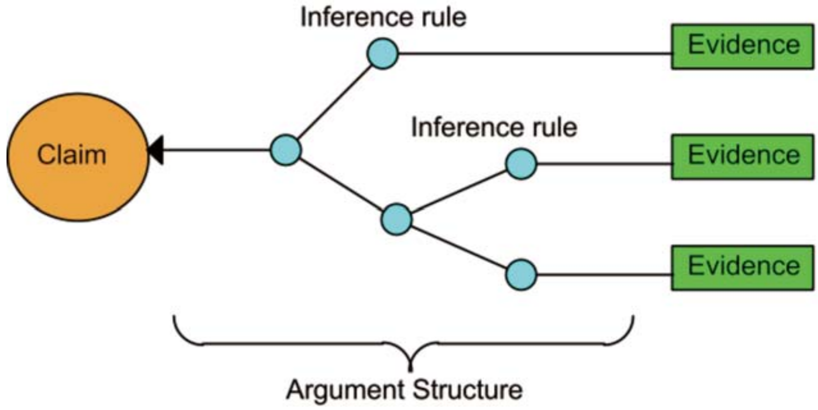
FEBRUARY 2010



Source: Walker (2005)

Figure 4-2 Supply Chain

Safety cases



- "a documented body of evidence that provides a convincing and valid argument that a system is adequately safe for a given application in a given environment"

Elements of a "Case"

- Claim about a property of the system or some subsystem, with some confidence.
- Evidence that used as the basis of the trust argument. This can be either facts (e.g. based on established scientific principles and prior research), assumptions, or sub-claims, derived from a lower-level sub-argument.
- Argument linking the evidence to the claim, which can be deterministic, probabilistic or qualitative.

Types of argument

Deterministic or analytical application of predetermined rules to derive a true/false claim (given some initial assumptions), e.g. formal proof (compliance to specification, safety property), execution time analysis, exhaustive test, single fault criterion

Probabilistic quantitative statistical reasoning, to establish a numerical level, e.g. MTTF, MTTR, reliability testing

Qualitative compliance with rules that may have an indirect link the desired attributes, e.g. compliance with QMS and safety standards, staff skills and experience

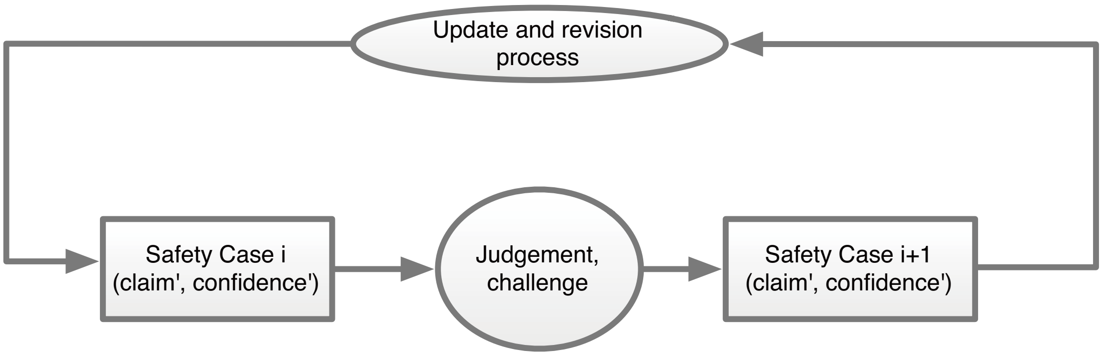
Making arguments explicit a key idea
Separating evidence from information

Communication and reasoning

- Structured safety and assurance cases have two essential roles:
  - communication is an essential function of the case, from this we can build confidence
    - boundary objects that record the shared understanding between the different stakeholders
  - a method for reasoning about dependability (safety, security, reliability, resilience ...) properties of the system
- Both are required to have systems that are trusted and trustworthy

Safety case process – building confidence, challenging assumptions

- Captured in safety management system and in meta-case
- Challenge and response cycle essential
- Proof as a social, technical, adversarial process

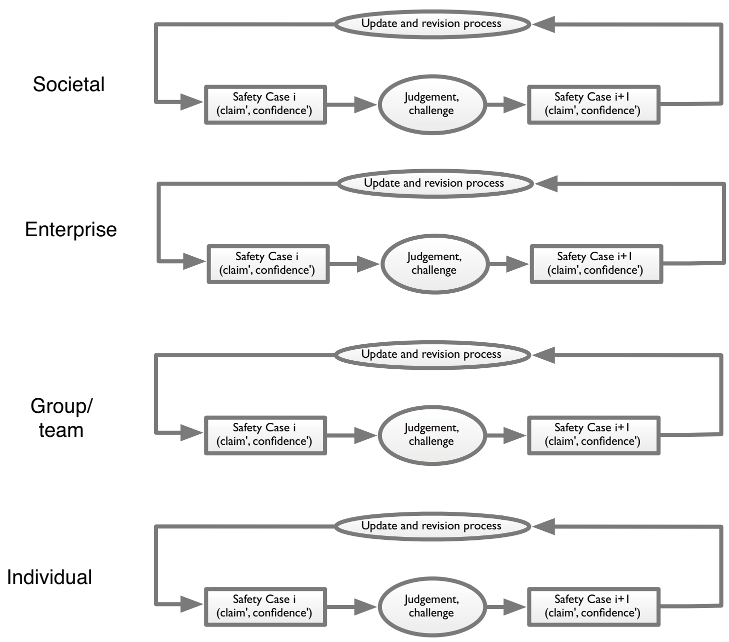


CSR Building confidence in a computerised world
www.csr.city.ac.uk

Adelard

Safety case process – building confidence, challenging assumptions

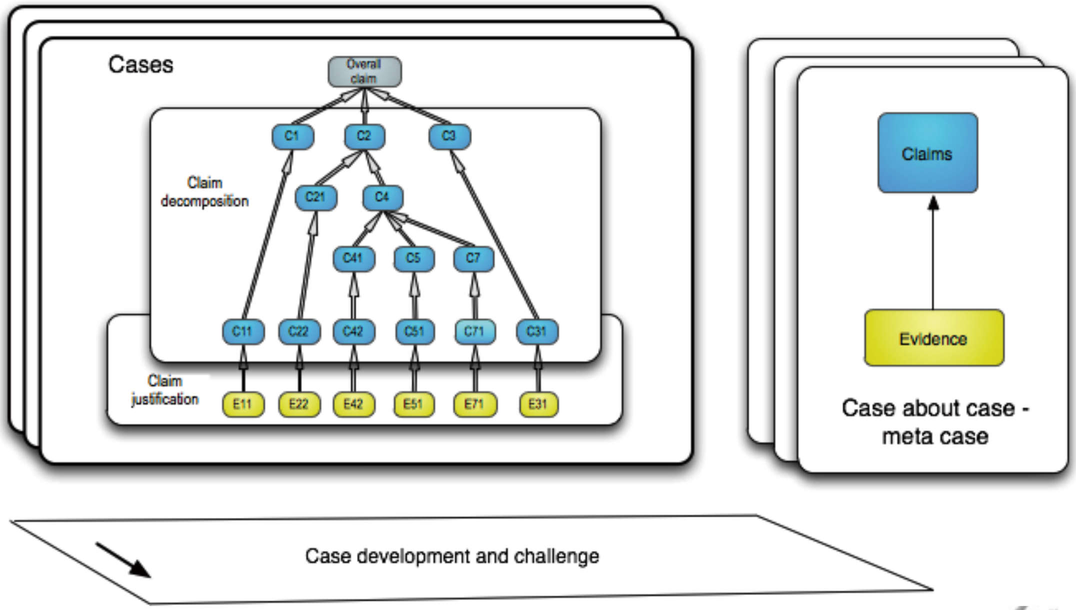
- Captured in safety management system and in meta-case
- Challenge and response cycle essential
- Proof as a social, technical, adversarial process



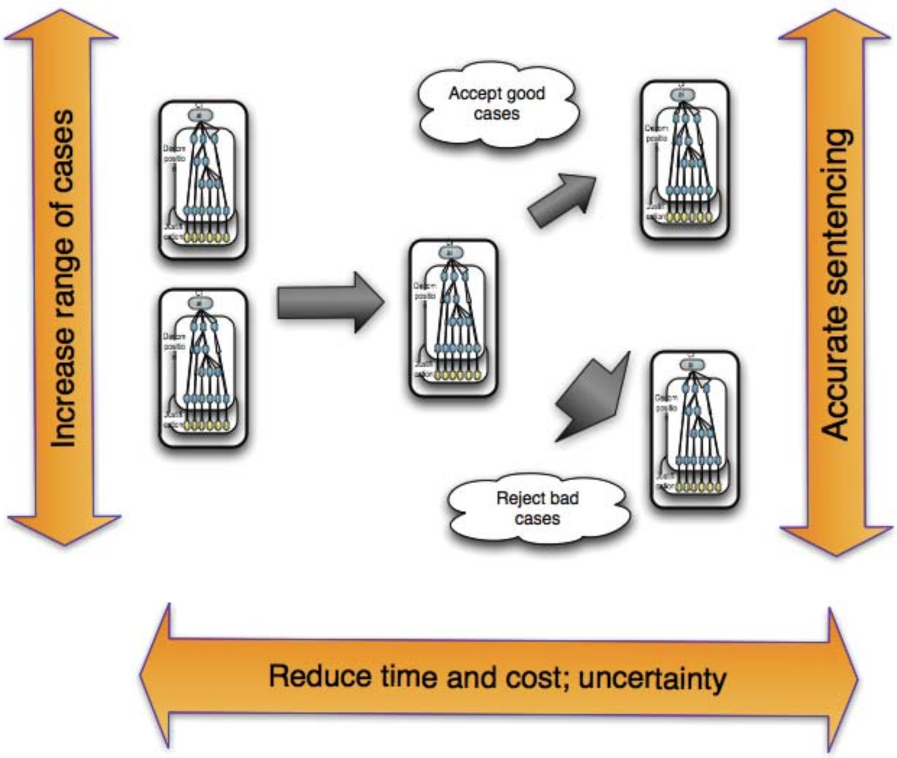
CSR Building confidence in a computerised world
www.csr.city.ac.uk

Adelard

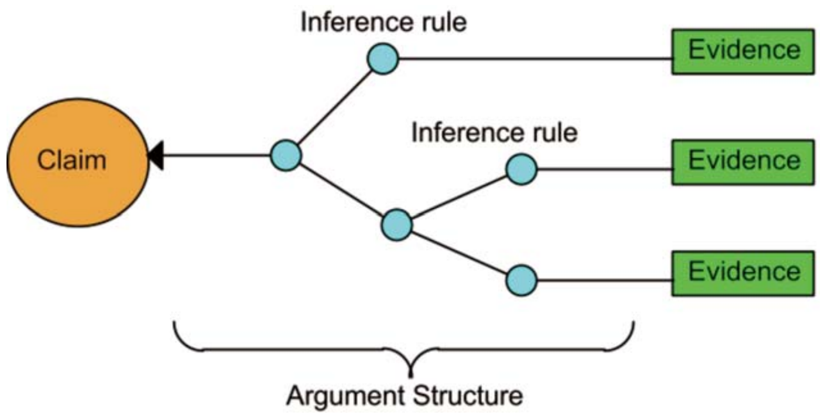
Reasoning, communication, confidence



Objectives

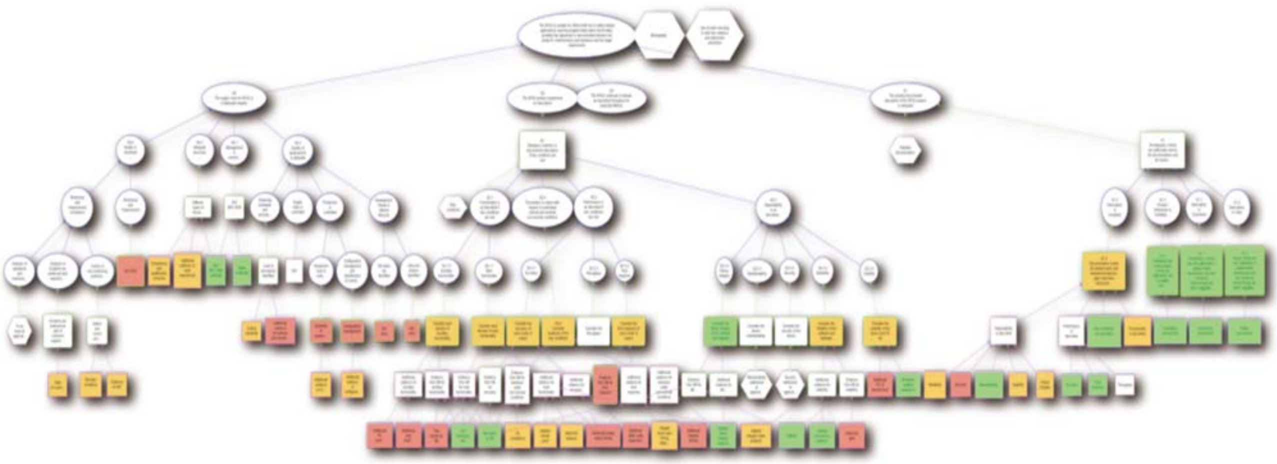


In theory ...

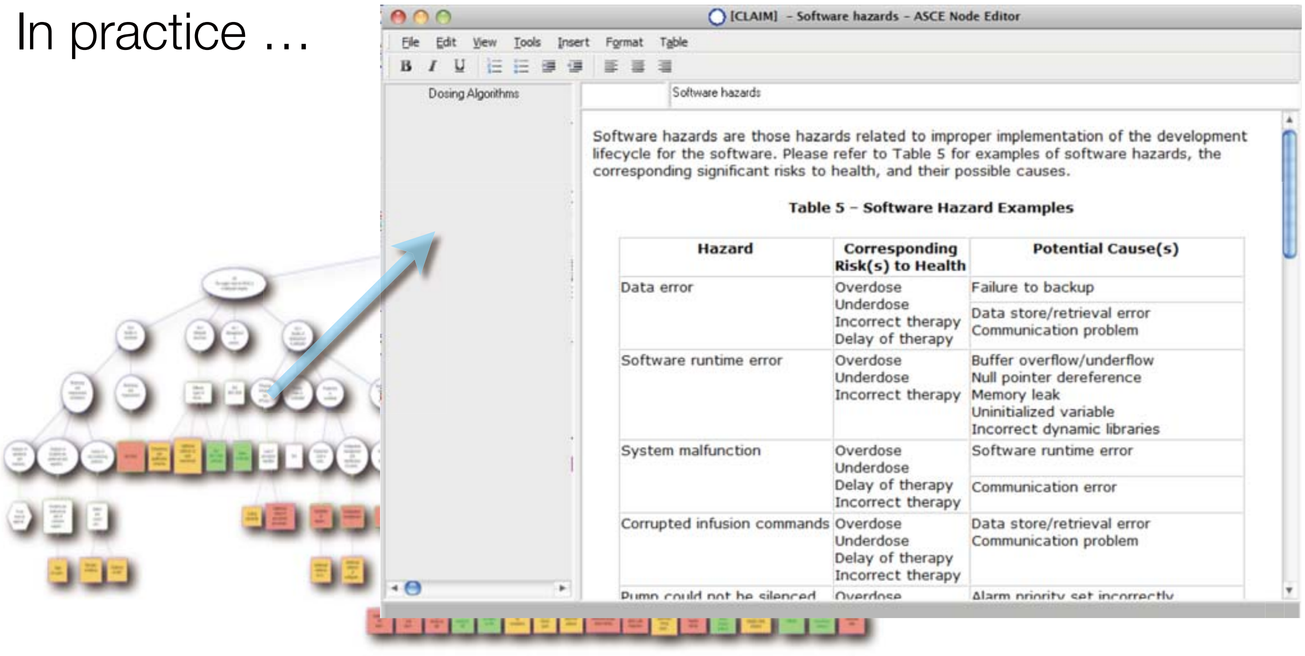


- "a documented body of evidence that provides a convincing and valid argument that a system is adequately safe for a given application in a given environment"

In practice …



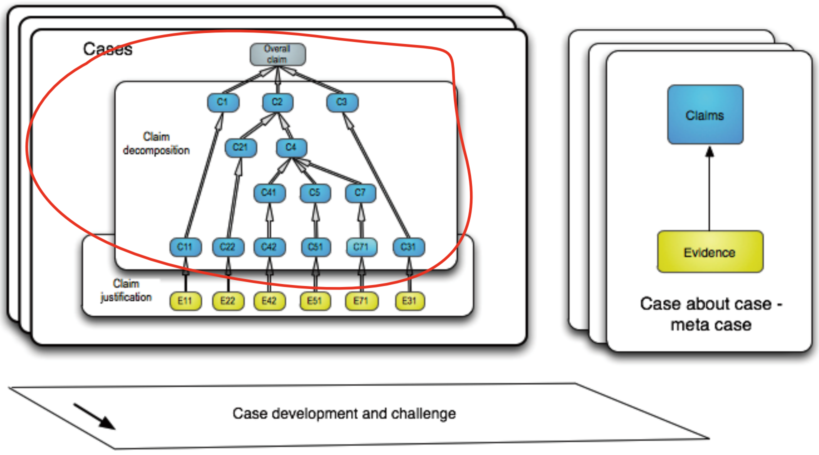
In practice …



Architecting claim structure

Claim structure

- creative strategies
- claims language
- templates



CSR Building confidence in a computerised world
www.csr.city.ac.uk

Adelard

Approaches

CSR Building confidence in a computerised world
www.csr.city.ac.uk

Adelard

Cases - argument styles

We have done what we were told to do (a *standards compliance* argument)

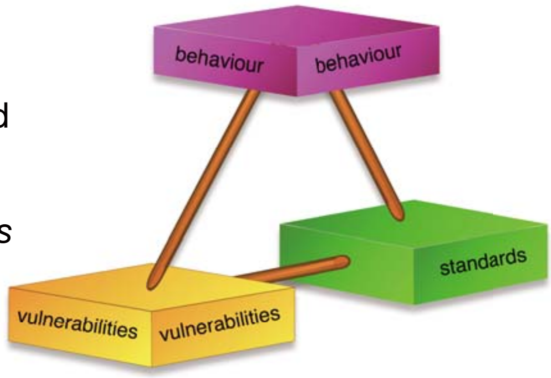
The system achieves the behaviour required (*safety properties* satisfied)

The system does not do bad things (*hazards addressed, vulnerabilities mitigated*)

Also

We have tried very hard (a *process argument*) to achieve dependability

Often a mixture of styles will be incorporated into a single case.

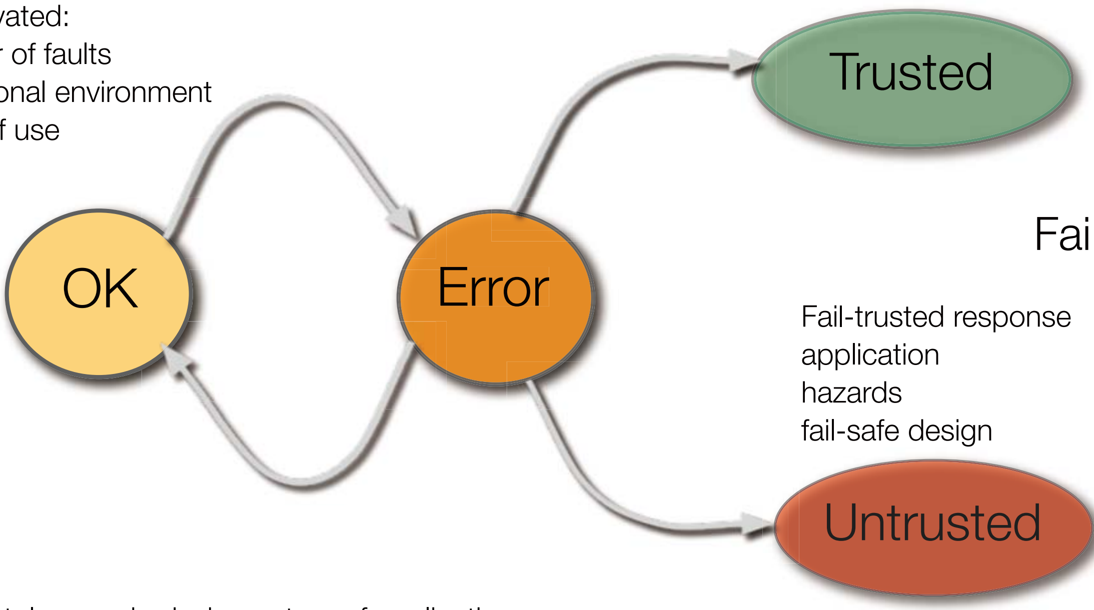


Standards and regulations

- Important part of case
- Can play different roles
  - Which needs to be justified
- But issues of validation
  - process -> product
  - techniques -> SIL achieved
- Need to innovate
  - Technology development V&V moves on
  - Use of COTS products
  - Product lines
  - Compliance can be expensive

Assurance strategies - behaviour

Fault activated:
- Number of faults
- Operational environment
- Mode of use



Failures

Fail-trusted response
application
hazards
fail-safe design

fault tolerance in design nature of application --
self healing, grace time

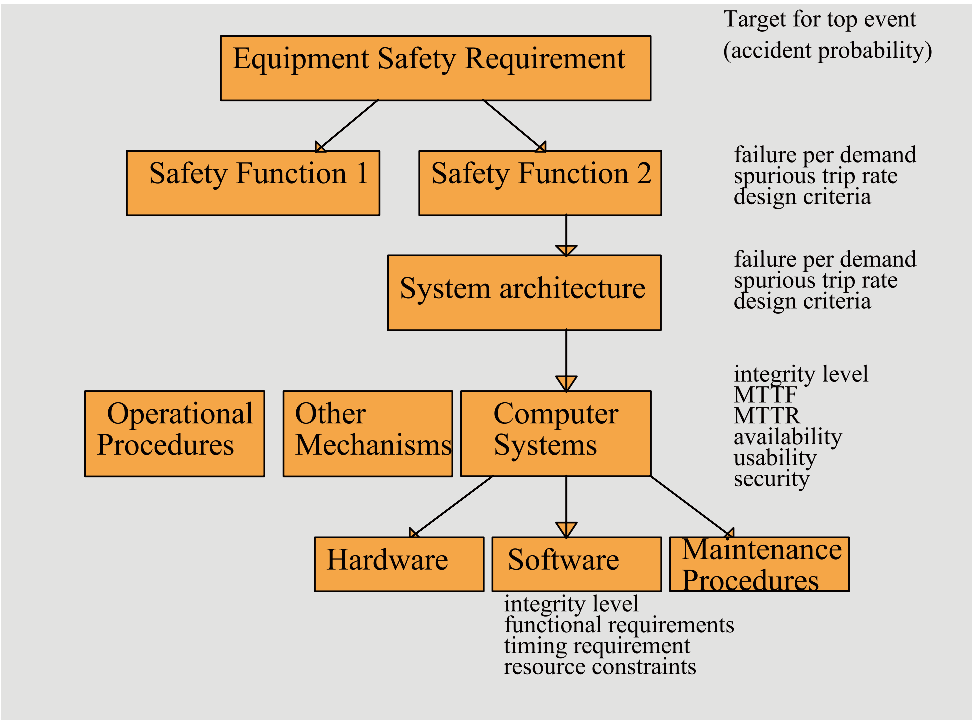
Strategies on behaviour

- Strategy – N No critical/significant fault or unsafe feature exists (the beast has no teeth, claws)
- Strategy –W Wrapper/containment argument – no failure or feature of the component can lead to hazard (the beast is in the cage)
- Strategy –R Restoration argument – any failure can be detected and recovered from (the beast can always be put back in the cage)
- And probabilistic variants of these

Safety properties and claims

- System safety analysis identifies hazards; these are amalgamated and abstracted into safety properties.
- Safety properties can be functions (shut down when T> 500), invariants (min sep always >2 miles) or purely descriptive (competency and culture).
- For each safety property address all attributes to increase completeness.
- As the design progresses need to consider derived properties arising from hazards introduced by the implementation.
- Non-functional system properties evolve
- May be claim limits

Architecture and functional claim expansion



Claim attribute expansion

- Claims can be broken down into claims about different attributes for the various sub-systems, e.g.:

reliability and availability
usability (by the operator)
security (external attack)
fail-safe response
functional correctness
accuracy
time response
robustness to overload
maintainability
modifiability, etc.

Restricted types of claim expansion

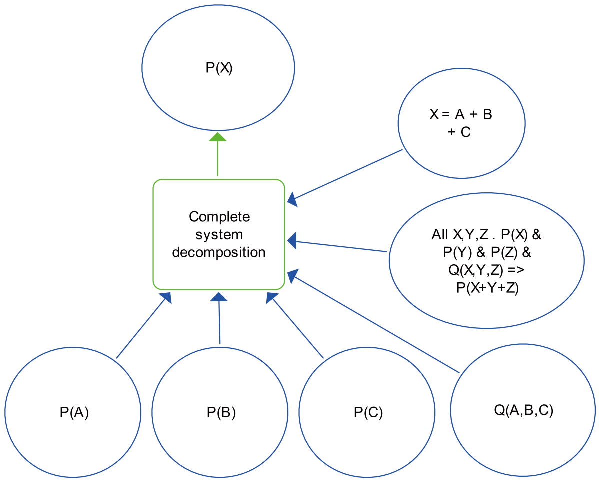
- Claim expansion language initially unconstrained
  - CAE
  - (also of course GSN)
- Empirically found a small set of constructs useful
- These enable more formal underpinnings and pragmatic checklists and tables
- Uniformity and regularity in cases
- Gradually introduced in our work
  - Part of work for the nuclear industry

Main types – keywords	Comment
architecture	splitting a component into several others
functional	
property decomposition	splitting a property into several others e.g. set of attributes
infinite set	inductive partitioning (e.g., over time)
complete	capturing the full set of values for risks, requirements, etc.
monotonic	the new system only improves on the old system
concretion	making informal statements less vague
generalises	property shown for one member of a class and generalised to all others
an-instance-of	properties shown for all components of a certain class

Pattern hierarchy and graphical summary

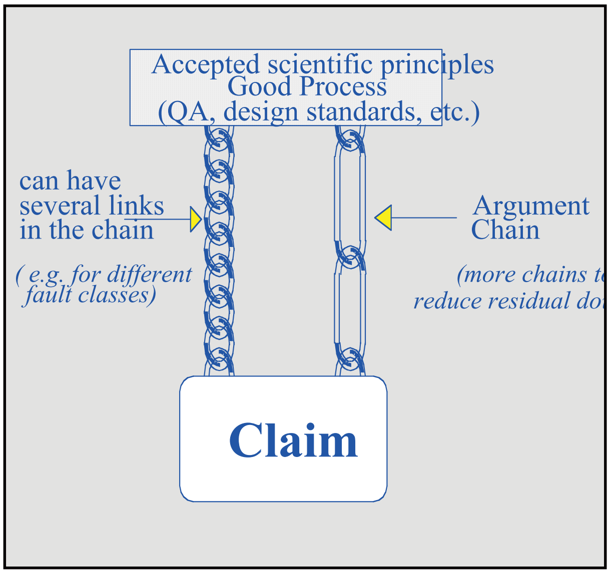
Partitioning decomposition

- Derive checklists for claim decompositions based on the formal work
- Once the structure is understood, the checklists are a way of verifying the structure is correct
- The checklists are informal but provide a route for more rigour if necessary



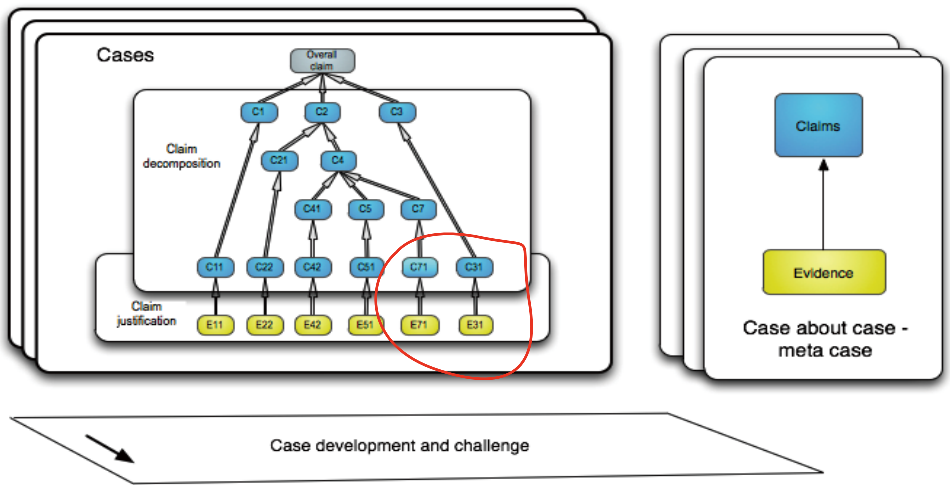
Argument metaphors

- Architecture of cases
- There is a parallel between architecture and argument structure
- e.g. in use of diversity, single failure criterion, sensitivity studies
- metaphors of "belt and braces", "legs to stand on"
- formalisation difficult and current research topic



Map evidence to claims

- iterative selection of techniques that generate evidence



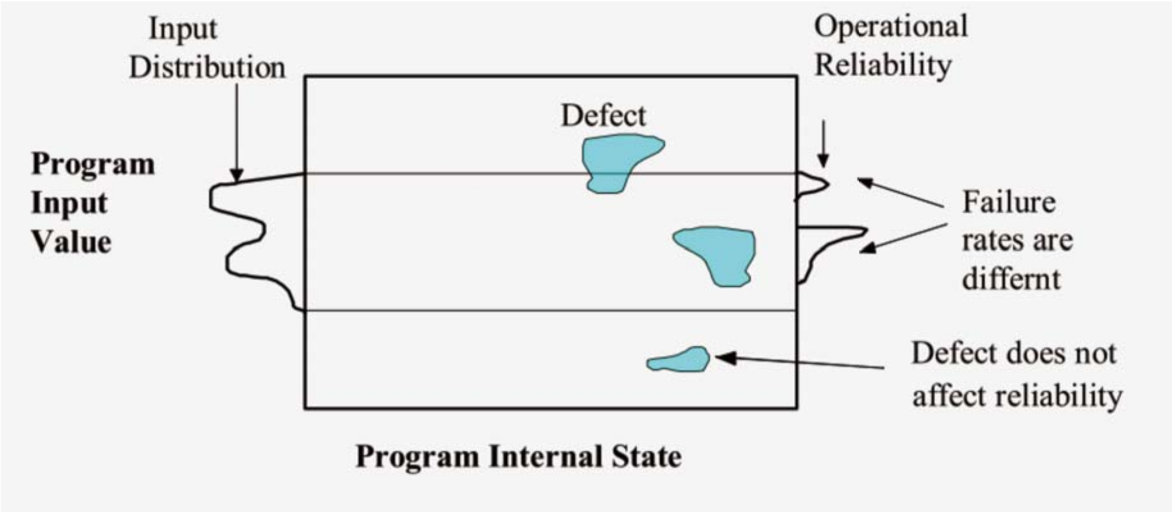
Selecting techniques and activities to generate evidence

- Catalogues of techniques e.g. in IEC 61508 Part3
  - P Bishop book
- Standards leave it as “exercise for the reader” in justifying selection
  - Supported by case
- Two useful mappings are
  - Activities/techniques → role in case
  - Attributes -> techniques
- Examples tables

Technique	Aim	Category	Assurance achieved	Effort	Expertise
Competence management	Assess competency management. Improve software quality by team with adequate competence.	FP	Indirect assurance from competence of development team.	Some additional management overheads.	Low, although assessment of requirements needs domain knowledge
Review of requirements process	Assess requirements process and requirements traceability.	FP	Increase confidence in requirements validity and satisfaction.	Information gathering may take a long time, depending on the complexity of the system.	High, as it needs to focus on what it is important. Need understanding of the system, vulnerabilities, weaknesses in both documents, process and specification
Review of quality of supply					
Supplier competency	Improve software quality by team with adequate competence.	FP	Indirect assurance from quality of development process.	Low	Low.

Reliability and process models

The software failure process

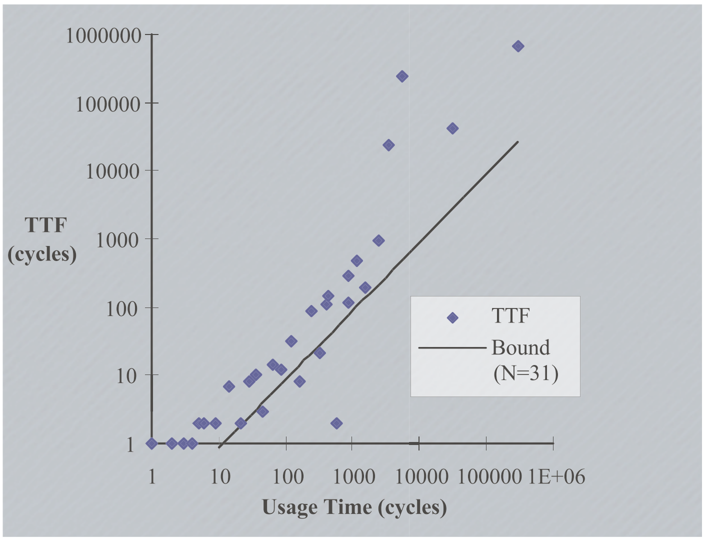


- stochastic nature from sampling input space
- “paradox” of deterministic yet stochastic in behaviour

Conservative long term prediction

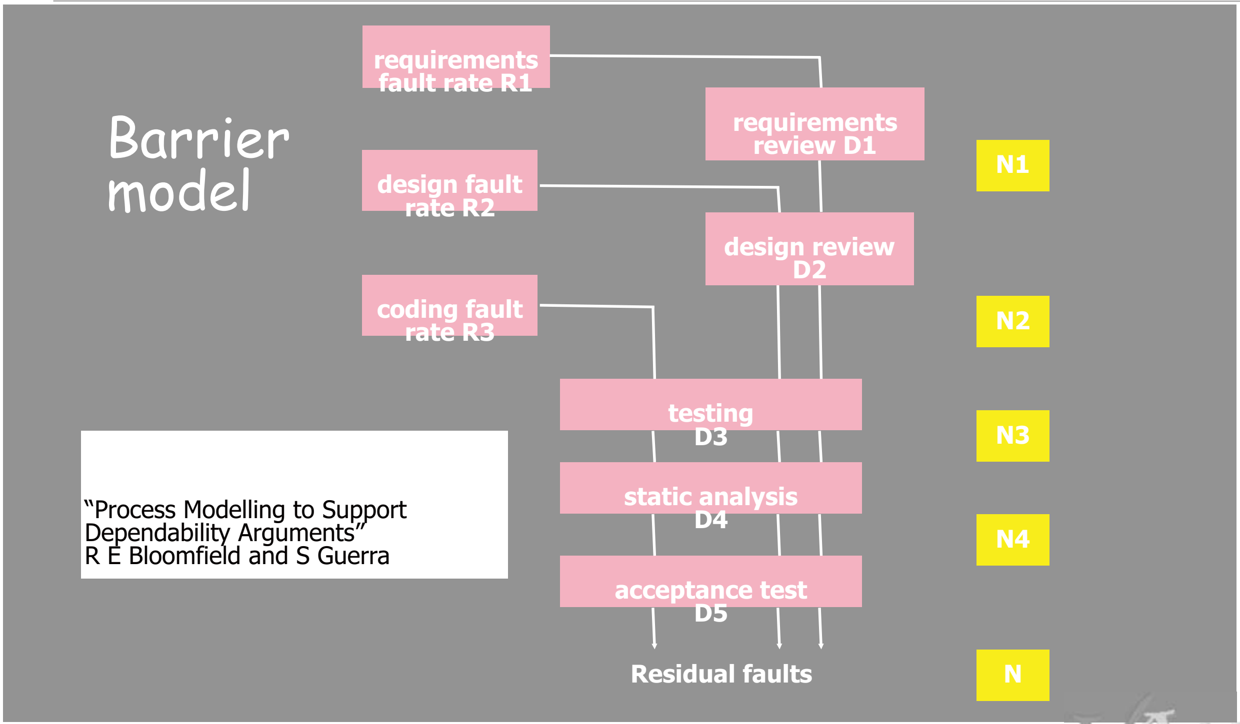
$$MTTF_T > e.T / N.d$$

Confirms every engineers intuition

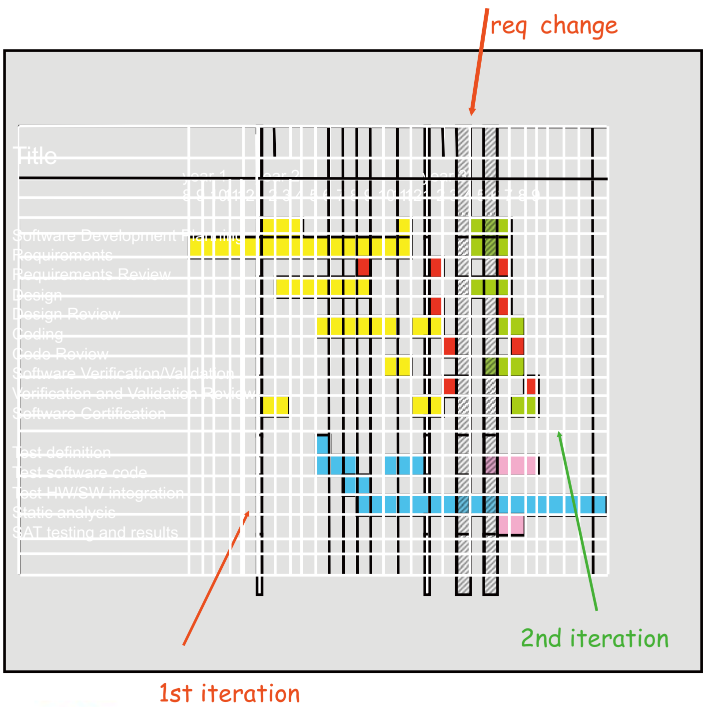


P.G. Bishop and R.E. Bloomfield, ÒA Conservative Theory for Long-Term Reliability Growth PredictionÓ. IEEE Trans. Reliability, vol. 45, no. 4, Dec. 96

Software development process



Use the results of the modelling

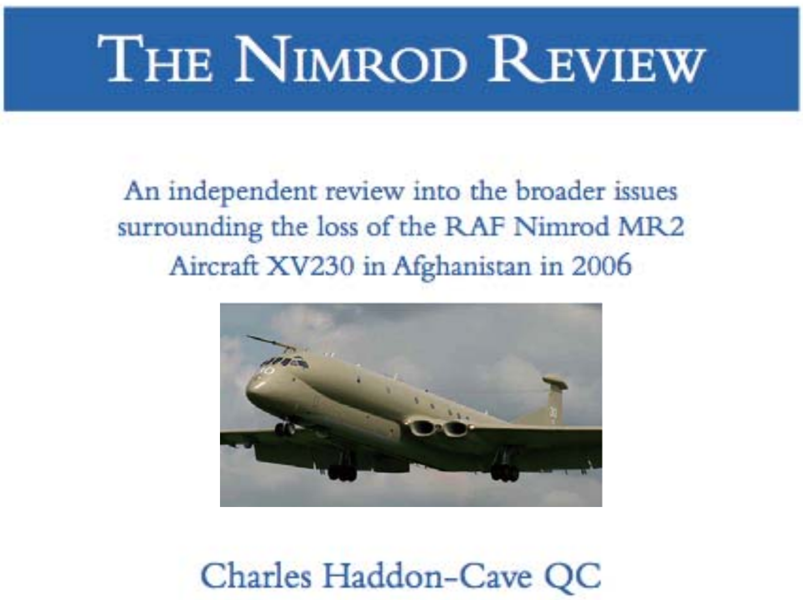


- Estimate residual faults.
- Reliability prediction techniques.
- Identification of weak areas in the process.
- Aiding process improvement
- Explore hypothesis as:
  - "what happens if design fault detection is increased to 90% by the use of tool xyz?"

Is this enough?

- If we have a claim decomposition that we think is adequate
- Is this enough?

Can we trust evidence?


THE NIMROD REVIEW

An independent review into the broader issues surrounding the loss of the RAF Nimrod MR2 Aircraft XV230 in Afghanistan in 2006

Charles Haddon-Cave QC

Research and development landscape

Research and development

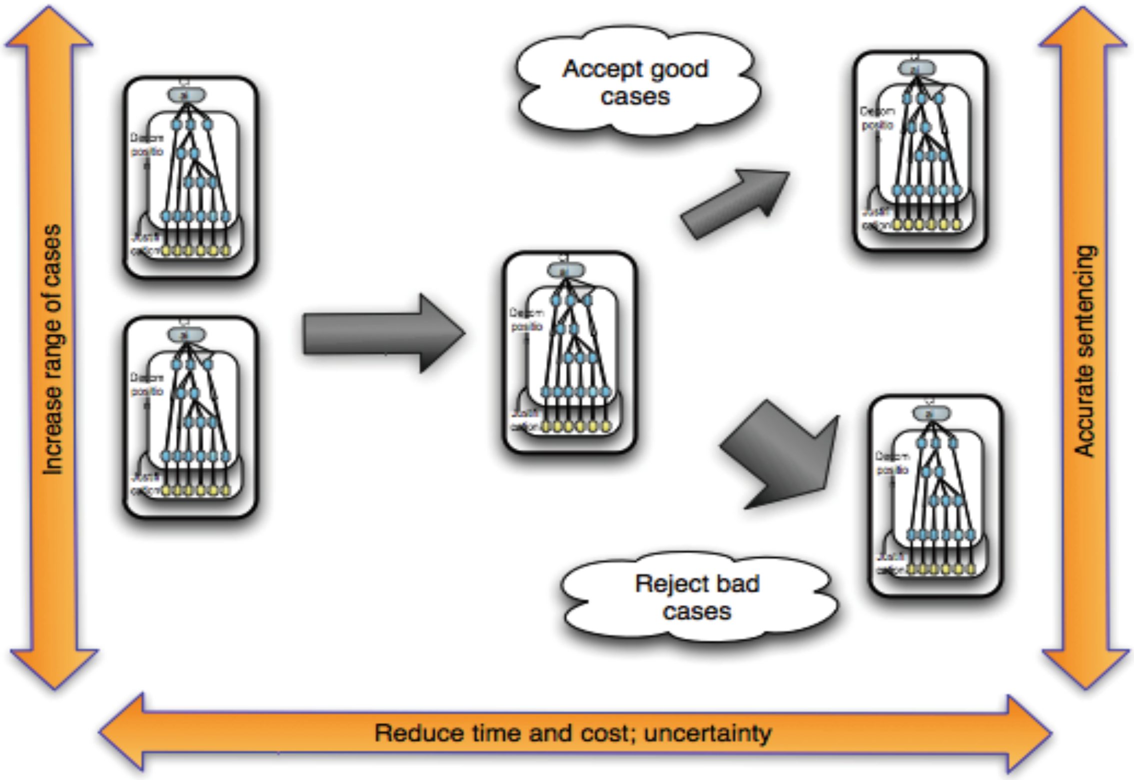
- Structures and scope of cases
  - How to justify the structure
  - Use of formal structures
  - Structures for different types of COTS components
  - Compositionality
  - Socio-technical perspective
  - Security, resilience and other cases
- Risk communication and scalability
- Role of standards
  - How to integrate standard compliance arguments
- Model based System/hazard analysis

- Styles of cases
  - Black-box
  - LowSIL
- Systems and cases
  - Architectures
  - Diversity
- Stopping rules
  - Claim limits and justification of numerical claims
- Confidence
- Evidence generation
  - Techniques and software analysis
  - Focused proof
  - Combing static/dynamic

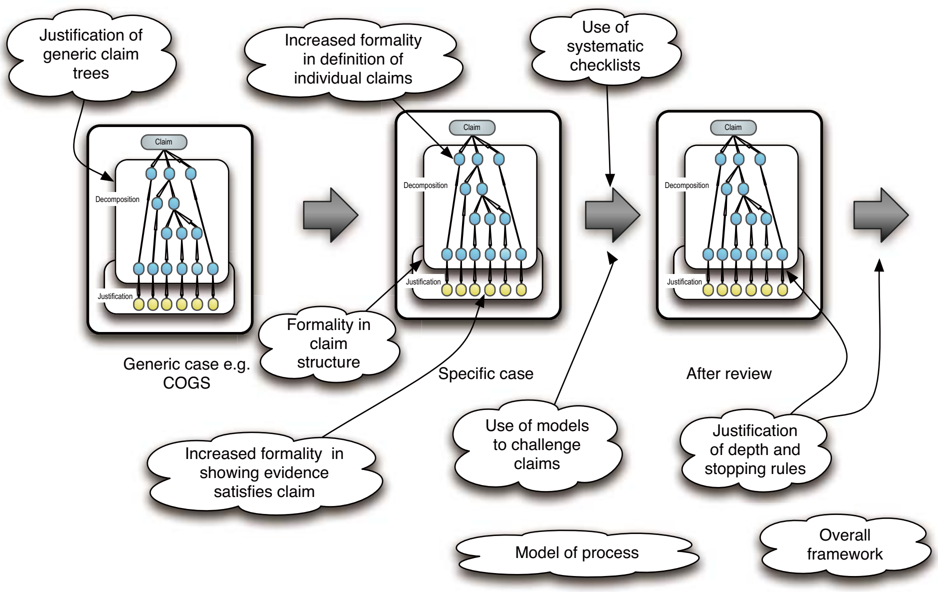
CSR Building confidence in a computerised world
www.csr.city.ac.uk

Adelard

Some drivers for research



Role of formality



Confidence

Aleatory and epistemic

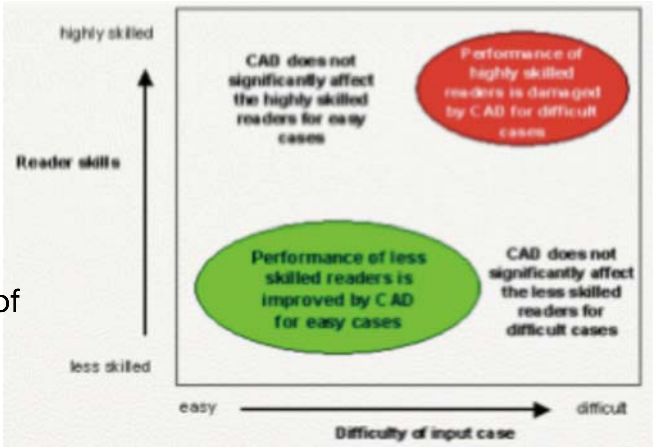
Work on confidence - summary

- Interpret existing practice in terms of confidence
  - Nuclear SAPS, ACARP in SOUP and SOCS report, CAA Regulatory oversight
- Empirical short study on assessors and SIL judgements
- Modelling of confidence in SILS, show impact, concepts and make speculative advice on standards.
- Confidence and legs (Littlewood, Bloomfield DSN)
- Extensive analysis of simple BBNs (Littlewood and Wright)
- Theoretical work on conservative approach, and later more useful bounds (TSE)
- Aleatory and epistemic distinction and dealing with system architecture/argument structures (Littlewood and Rushby)
- Threat models
- Stress claim/confidence pairs

Socio-technical

- A socio-technical perspective on assurance cases:
  - In addition to claims that physical hazards, security threats have been addressed
  - Define a range of vulnerabilities (narrow scope, misaligned responsibilities, undifferentiated users, adaptation, automation biases, non-independence of arguments) and develop arguments of how they might be addressed.
  - Develop methods for review wrt socio-technical issues

Ideas taken from EPSRC INDEED and DIRC projects



Scale and complexity

- Assurance Cases scale
  - Claims, Arguments, Evidence in Generic Design Assessment (GDA) New Nuclear Build
  - FDA and infusion pumps
  - Defence Systems
  - Global component manufacturer
  - Financial processing system
  - ASCE user base for structured assurance (dependability) cases

Scaling pragmatics

- Pragmatics
  - CAE leads to focus (cf compliance cases)
  - abstraction, modularity, timebands
  - assumption and knowledge engineering - pragmatics
  - reference out to other documentation, cases e.g. for correctness
  - use notation of for what it is good for
  - guidance, templates, capturing best practice and domain specific regulations
  - limit graphical wallpaper
- issues
  - systems of systems
  - what to expose on interface, how to find relevant detail

Timebands

Characterisation1	*Threats/events*	*Example mitigations*
Generation 1-30 yrs	Obsolescence Organisation death/rebirth Major external events (economic, social)	Moore's Law, adaptation and evolution of the system as a whole See grid/group; long term risk analysis
Social time Months-Year	Staff turnover; relocation; restructuring, culture change	Training Change management
Processing cycle /Days	Procedure violations Equipment repair time	Redundancy in the system; diversity; compliance management
Problem solving /Hours	Problems with master records, assessing problem failure	Part of normal operation. Embedded in overall system design.
Cognitive /Seconds	Distractions, slips/lapses	Either reduced by equipment reliability and checking or caught at problem solving level.
Biological/Equipment2 <0.1s	Equipment component failure	Machine based checks Fault tolerance

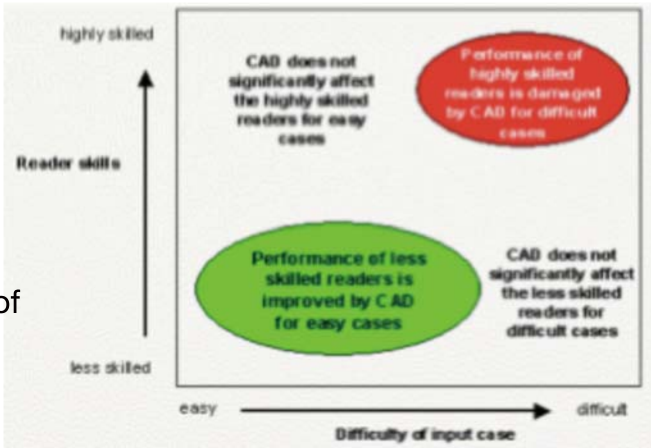
Dynamic cases

- claim structure more static;
  - includes claims about ability to update and respond
  - as pattern for a range of scenarios
    - adjust, update, select
    - assets change
    - need to make rely assumptions clearer (e.g. positive behaviours)
  - pattern for different parts of resilience curve
    - normal levels of threat and response
    - incident response
    - heightened threat levels

Socio-technical

- A socio-technical perspective on assurance cases:
  - In addition to claims that physical hazards, security threats have been addressed
  - Define a range of vulnerabilities (narrow scope, misaligned responsibilities, undifferentiated users, adaptation, automation biases, non-independence of arguments) and develop arguments of how they might be addressed.
  - Develop methods for review wrt socio-technical issues

Ideas taken from EPSRC INDEED and DIRC projects



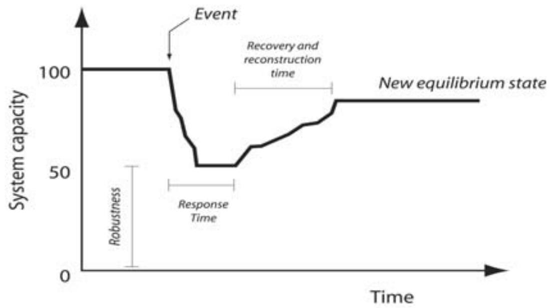
Scale and complexity - some challenges

- Organisational and confidentiality boundaries
- Dynamic cases
- The importance of the socio-technical
- The importance of detail and possible limits to abstraction
- Interested in risks from systems
  - non-linearities, cascades, adaptation, emergent properties....
  - need to extrapolate.. theories.
- Need to develop with a “fusion” of complexity science, risk analysis and computer science
  - find the right combination of concepts, analysis, data, models and theories

CSR Building confidence in a computerised world
www.csr.city.ac.uk

Adelard

Concepts - resilience viewpoint



- *Type 1*: Resilience to design basis threats. This could be expressed in the usual terms of availability, robustness, etc. It could be bounded by credible worst case scenario.
- *Type 2*: Resilience to beyond design basis threats. This might be split into those known threats that are considered incredible or ignored for some reason and other threats that are unknowns.
- Attacks on intangibles - these are also societal assets, not just CIP
- Does addressing Type 2 help with Type 1?

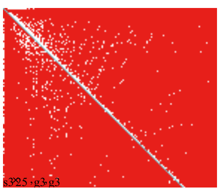
CSR Building confidence in a computerised world
www.csr.city.ac.uk

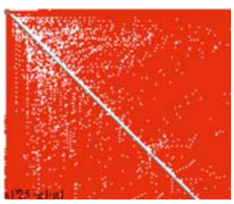
Adelard

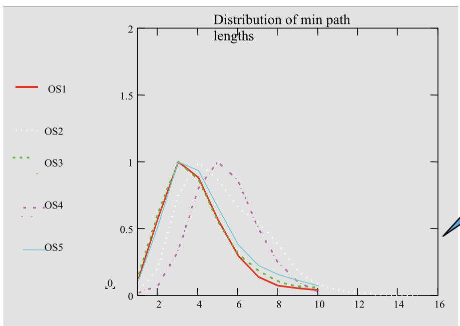
Complex systems

- common mode and cascade failures
  - extrapolate from small losses - complex systems models
    - preferential attachment, highly optimised tolerance and self organised criticality
  - COTS software
  - critical infrastructure modelling - interdependencies
- small changes
  - method for evaluating changes to complex, evolving systems
  - does this small change have a small impact? regulatory risk
- issues of experimental methodology

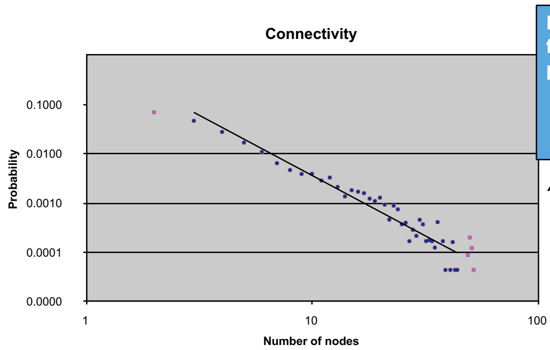
Sparse, long range, weak clustering







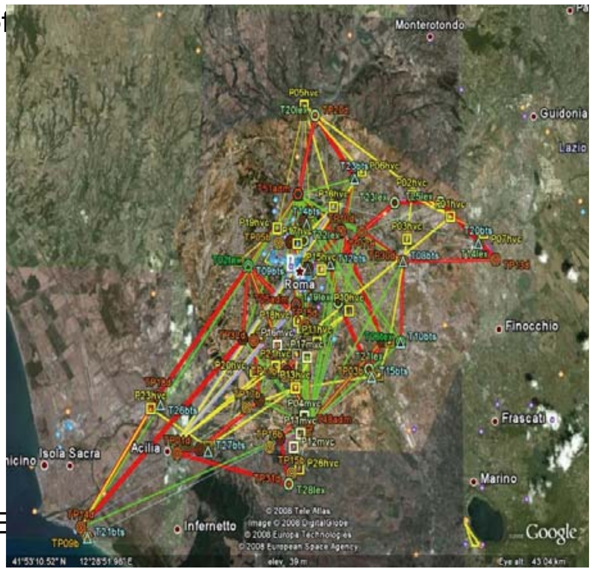
Similar distribution of min path length



Evidence of scale free network - power law

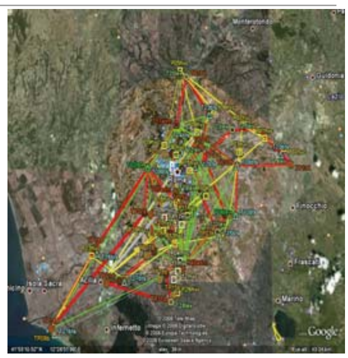
Rome Scenario implementation

- Service layer, Physical layer;
  - Same formalism used to model at both levels o
  - What do we gain/lose with increased detail?
- Nodes and Physical/Abstract Links;
  - HVC, GSMTrunk;
- Dependencies;
  - PhonetoMVC;
  - DC Power-Flow calculations (ETHZ);
- Boundaries;
  - Power supply to Telco exchanges;
- Parameter values;
  - Failure rates, Repair rates;
- Characteristics of Nodes and Physical Links;
  - Link capacities, voltage levels, line resistance (E
- About 500 nodes
- Issues of research methodology, testbeds, scaling, realism

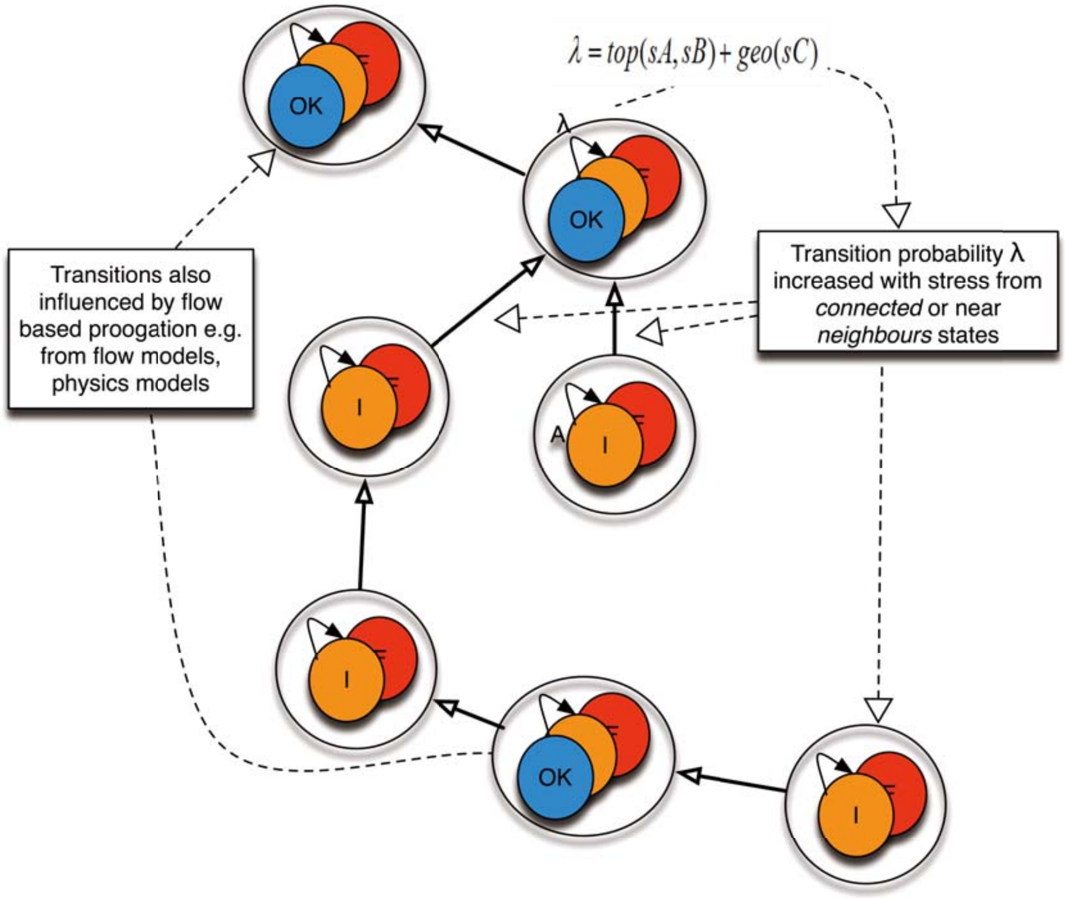


PrIA models

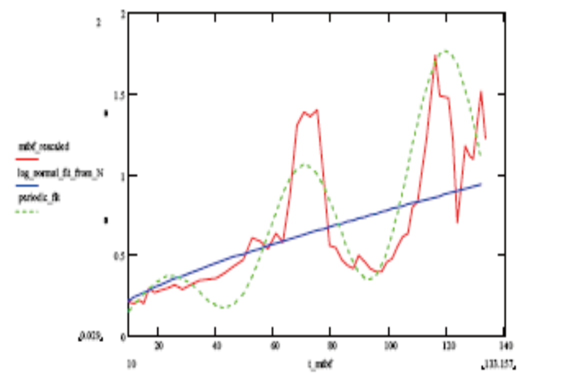
- used SANs (stochastic activity networks) and Mobius Modelling Tool to define parameterised continuous time Markov models
- finite state atomic component that mutually interact to make impairment and failure "contagious":
  - rates of transition to impaired and failed states are functions of the states of nearby components (stress).
- embedded deterministic sub-models that can relate the "dynamics" of some subsets of the components in other specified ways
  - e.g. DC approximate power flow model for power flow components
  - e.g. telco service model.



Stochastic associations - sources of dependency and cascades



CI dependencies



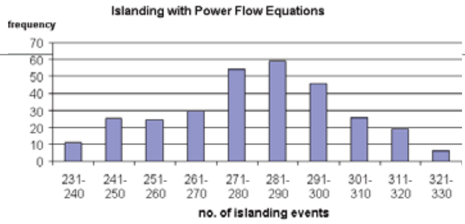
- shared services or functions
- shared resources
- similar policies
- similar assets attracting correlated attacks
- similar components (e.g. COTS)
- traffic/load dependencies
- common environmental effects (flood, fire, disease)
- poisoning and spreading of failures
  - (e.g. by traffic on a telco network, denial of service by device failing on network and causing flooding of network)
- human networks e.g. maintenance teams,

These can lead to a combination of unanticipated connectivity, greater impact of failure, and faster, cascade events.

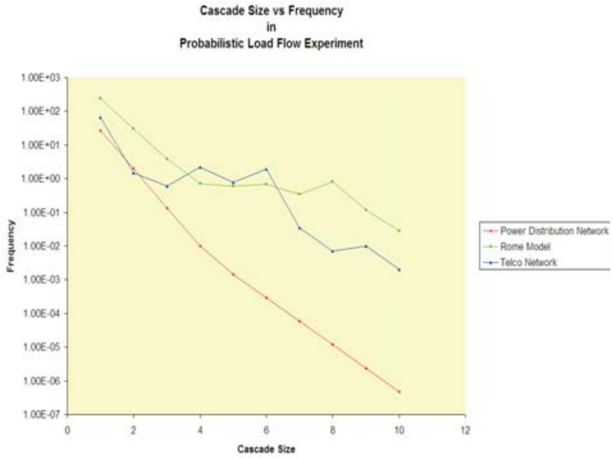
CSR Building confidence in a computerised world
www.csr.city.ac.uk

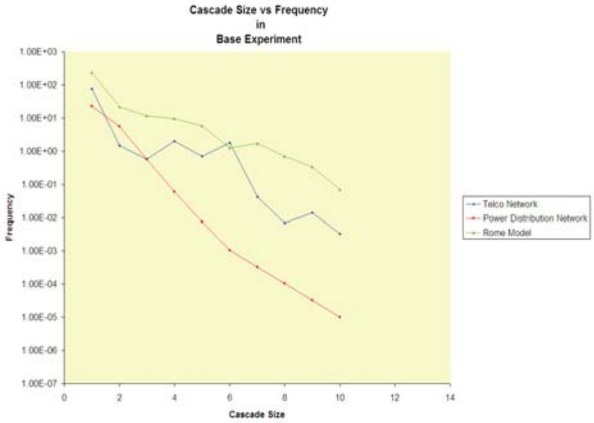
Adelard

Typical results



Size of loss





Conclusions

- Wide experience with structured Safety and Assurance Cases
  - threat and promise
- Claims, Arguments, Evidence provides a scalable framework
  - Adelard and public domain publications in 2011, give away
- Rich research landscape - claim structures, confidence, socio-technical vulnerabilities
- Open, complex, systems pose fascinating challenge
  - focused on resilience, interdependencies, cascades and change
  - issues of methodology
- Next steps
  - develop complex systems approach to cascade/rare losses in computer based socio-technical systems and interdependencies
    - investigate role of abstraction using Rome scenario

Conclusions

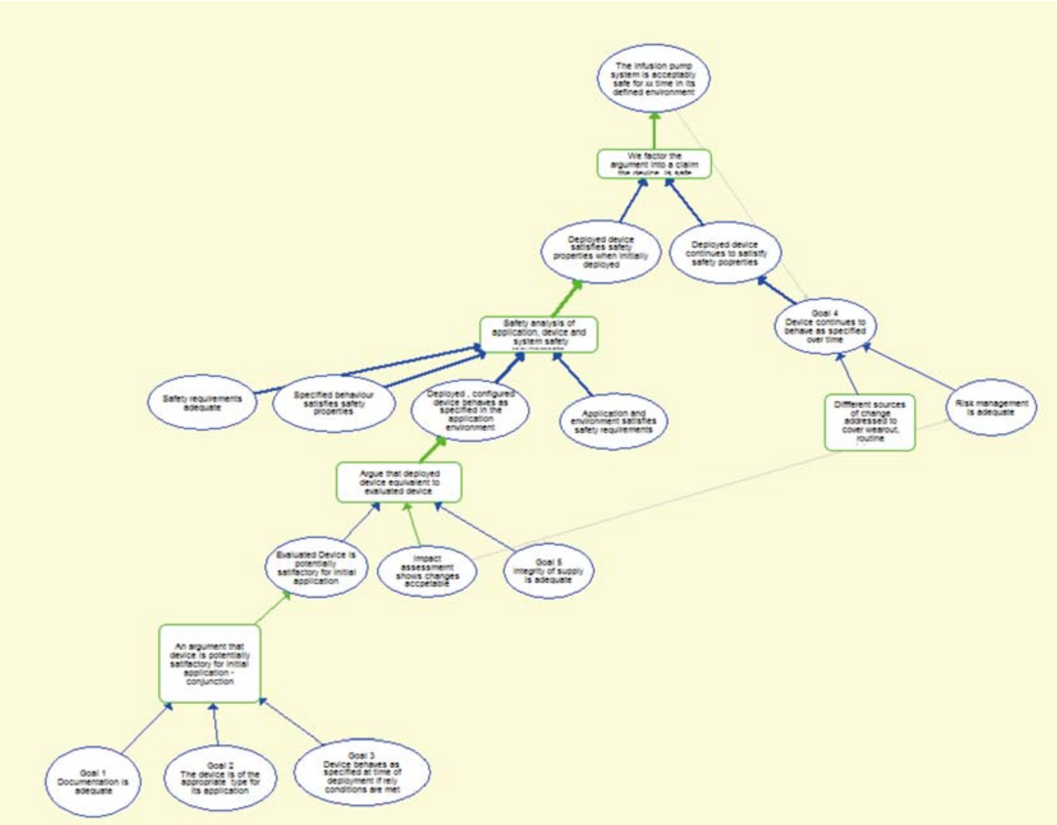
- Reviewed assurance case concept of claims, arguments, evidence - CAE
  - case, meta-case and confidence
  - Major strategies for architecting claim structures
  - Mappings between techniques and evidence
  - Technical approach for dynamic and static analyses
  - Supply chain experience from nuclear industry and financial services
- Extending notion into resilience and assurance cases and SCRM
- Aspiration to consolidate, publish and give away

Acknowledgments

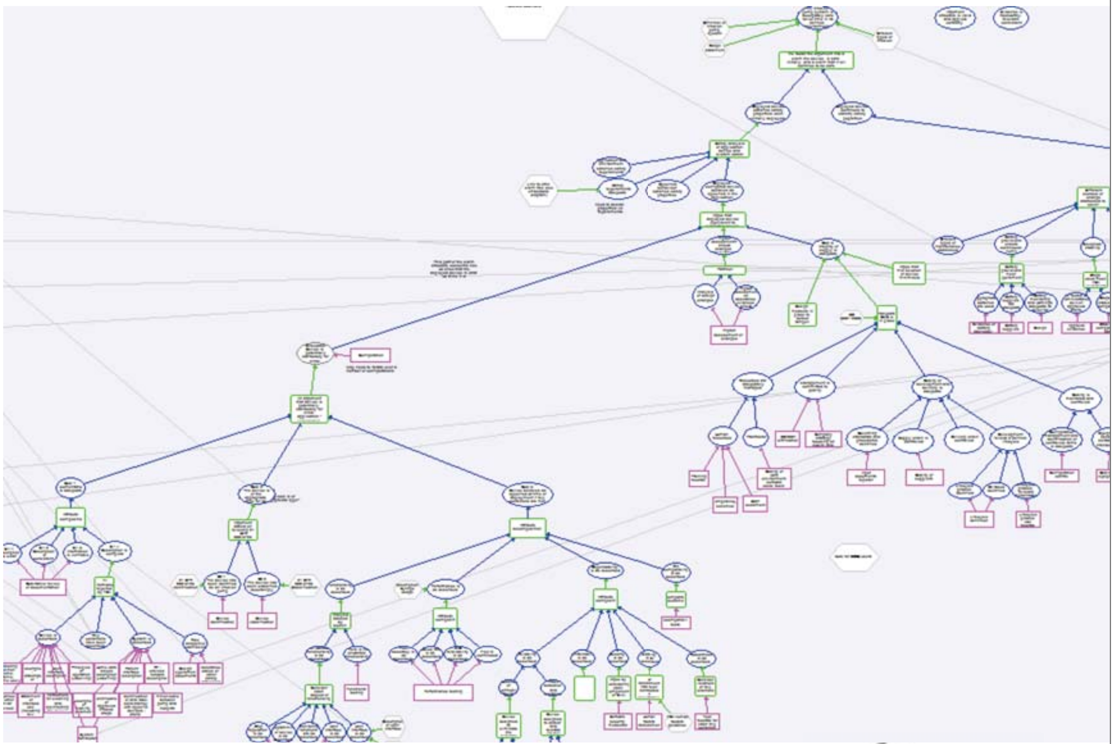
- Colleagues in CSR and Adelard, particularly Peter Bishop, George Cleland, Lukasz Cyra, Sofia Guerra, Dan Sheridan, Bev Littlewood, Andrey Povyakalo, Lorenzo Strigini and others

Additional material

Example



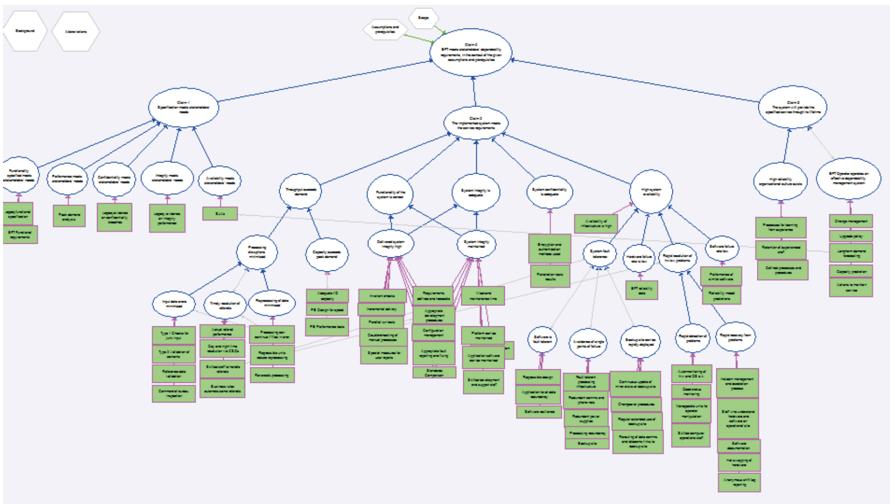
Example



Financial services dependability

High volume

- Socio-technical perspective
- Deployment decision
- Range of stakeholders



Meta-case

Structuring judgements

- We can show from a fairly rigorous model and conservative assumptions that we need four parts to a case:
  - A judgement on the safety given the context and argumentation and evidence
  - A judgement on the argument structure, application and claim decomposition and backing evidence
  - A judgement about the quality of the evidence
  - A judgement of the context

Meta case

- Case history and development process. This would describe the development of the case, the judgements made about it, the history of challenge and confidence building that has been done. This would document the safety case process
- Challenges and confidence building. This would describe and detail the challenges and confidence building measures. This would consider
  - diverse derivation of properties
  - use of different but claimed equivalent properties (e.g. best estimate timing, worst case)
  - use of different but related properties, different models, diverse tools
- It could be based on a Hazops-style keyword approach
- A risk based “red team” attack on a case looking for vulnerabilities based on experience (compare with preliminary hazard list, safety case fallacies) could also be applied.
-

Threat and promise

Maturity indicators

- ASCE statistics
- 250 organisations in 15 countries, many 1,000s users
  Key users:
  BAE SYSTEMS, QinetiQ, Boeing, Lockheed Martin, Raytheon, Thales, Westland, MBDA, General Dynamics, Northropp Grumann, AugustaWestland, Selex, Atkins, Quintec, Logica CMG, HVR, AWE
  Bosch, TRW, Moore Industries, Mira, Entec
  British Energy, BNFL, SKI, Framatome, AVN
  CAA, NATS, IAA, Eurocontrol, Indra, Advantage, CSE, Ebeni, Helios, Weston Aerospace
  Mitre Corp, FDA, NASA, Elekta Oncology, Cardinal Health, Medtronic
  Frazer Nash, Strachan and Henshaw, SSMG, NNC, ERA, Praxis
  Westinghouse, Ansaldo, Thales Rail, Network Rail
  MoD: Tornado, Harrier, Chinook, Jaguar, Puma Gazelle, JSF, Sea King, Merlin, ARC, U/water weapons, Helicopter Engines, ALM, PGB, Eurofighter/Typhoon, SUAV(E), Sub IPT, HMNBs Clyde & Portsmouth, Astute, TA, Bowman, DOSG, NW IPT, SSMO, LSSO, ARC, GBAD
- OMG standardisation, ISO 50126, Nato, FDA
- ... but need

The promise of assurance cases

- Innovation in systems and assurance technologies
  - Can see how to incorporate new evidence
  - Cope with change, principled non-compliance
- Innovation in justification arguments and evidence
- Expose lack of validation of standards, gaps in our knowledge
- Focus of assessment and challenge
  - Need supporting safety case process and meta-case
- Clarity in the basis for regulation and licensing
  - See shortcomings of present approaches
- Improved communication with stakeholders
- Improved knowledge management
- Scalable
  - From smart components to complex systems
- Multi-attribute
  - Dependability, safety , security

Threat of assurance cases

- Apply safety analysis to cases themselves to understand risks and mitigations
  - Systematically analyse the failure modes for safety cases, using a HAZOPS style technique
  - Rejecting satisfactory cases, accepting inadequate cases
- Expose lack of validation of standards, gaps in our knowledge
- Competencies and skills and deployment risks
  - need for more methodology, examples
- Negatives to avoid
  - outsourced, commoditised, lack of controlling mind
  - just another report - value marginalised, a cost
  - complex, unclear, inappropriate cases

Professor of System and Software Dependability
Director, Centre for Software Reliability
Founder Adelard LLP
reb@csr.city.ac.uk
College Building, City University, London EC1V 0HB
Tel: +44 20 7490 9450 (sec Adelard)
Tel: +44 20 7040 8420 (sec CSR)