Certification and evaluation -- current methods and future prospects

Robin E Bloomfield, CSR City University and Adelard

Background and state of practice

The certification that a system is fit for purpose, and continues to be fit for purpose as the environment, use and implementation change, is a complex socio-technical process. Historically in other sectors there has been a continuing move away from a standards, compliance based approach, to one where goals of system behaviour are justified. So called safety, assurance or dependability cases are becoming widespread and in many sectors (in Europe) are mandated by legislation.

I have been involved in the development of safety cases, standards and underpinning research for many years. My recent perspectives come from the wide user base of a pliant hypertext approach to presenting cases (the Adelard ASCE tool) coupled with our own application in the finance, air traffic management and nuclear sectors. A broad socio-technical view is provided by engagement with the UK Interdisciplinary Research Collaboration DIRC (www.dirc.org.uk).

The assurance case approach can be summed up by the slogan "its not how hard you have tried but the behaviour of the system that matters". For critical systems it is often the so-called non-functional requirements that are the hardest to justify. While extensive and detailed compliance with standards may be necessary it is not sufficient. Current standards are, on the whole, not validated and there are counter examples, especially for lower criticality systems (e.g. following IEC 61508 SIL2 recommendations does not necessarily produce SIL2 systems.

Our customary definition of a case is "a documented body of evidence that provides a convincing and valid argument that a system is adequately dependable for a given application in a given environment". This highlights some of socio-technical basis of trust and trustworthiness – who is being convinced? Should they be convinced? The definition can be unpacked further to drive an appraisal of current practice and an identification of future challenges: what is sufficient evidence? What does documented mean? What is dependable? What is adequate? How do we characterise the environment etc etc?

The "case" and associated supporting tools can be seen from a number of viewpoints

1. As a boundary objective between the different stakeholders who have to agree (or not) the claims being made about the system. To this end it has to be detailed and rigorous enough to effectively communicate the case and allow challenge and subsequent deepening of the case.

2. As an over arching argumentation framework that allows us to reason a formally as necessary about all the claims being made.

3. As part of a "signal processing system" – the licensing and certification process – that seeks to reject false claims and accept good ones. For critical systems the probability of false positives has to be very low (i.e. a very low chance of accepting a flawed system) and most regulatory approaches seem to achieve this. They may also have high false negative rate, rejecting acceptable systems.

Current best practice involves graphical summaries of the case, tracking of evidence status, propagation of changes through the case, automatic linking to requirements and management tools. However claim decomposition is normally very informal and argumentation is often not explicit. In fact the emphasis is on communication and knowledge management. The uptake of the approach is very strong in some sectors (e.g. most new defence aircraft in the UK are certified using this approach).). There is some empirical evaluation of the benefits of the "cases" approach: one academic study in the US and in the UK we have worked with the UK MoD assessing the benefits/costs of the safety defence standards.

Research challenges are everywhere: claims, arguments (confidence in fault freeness vs reliability?), evidence. At a user and societal level I see the need to scope the scale of the challenges

- To stand still - so that there are no disasters. The demands of current systems, complexity of new systems, time to market, pressures, the structure of the supply chain, competence, the scale of reuse, evolving threats all make this non-trivial and uncertain.

- To allow a choice of doctrine between conservative approaches to technology adoption and an ability to make informed decisions to exploit the benefits of using computer based systems. To effectively describe and communicate the levels of trustworthiness.

In well-engineered systems the limitations that we can convincing claim about dependability are likely to come from assumption doubt. We need to methods for dealing with this uncertainty and new approaches to common mode failure assessment. Some of my own research priorities can be seen within the recently awarded UK project Interdisciplinary Design and Evaluation of Dependability (INDEED) that I am director of. This has four main work packages:

Timing and Structure (led by Alan Burns, University of York). This will extend the time band model to serve as a basis for using time as a fundamental driver in structuring the architecture of dependable, socio-technical systems.

Adaptation and Diversity (led by Lorenzo Strigini, City University, London). This will deliver probabilistic models of diversity, adaptation and learning that allow designers to analyse how these influence system dependability and explore the consequences of design choices.

Responsibility and Trust (led by Ian Sommerville, St Andrews University). This will provide a means to understand, identify and mitigate potential failures in complex socio-technical systems arising from the responsibilities of the agents in the system and the trust relationships between them and to extend dependability cases to take account of this class of failures.

Confidence and Uncertainty (led by myself). This will deliver a method for structuring confidence-based dependability cases, based on a probabilistic interpretation of confidence. We will define a rigorous approach to the propagation of confidence in cases where claims are structured on the basis of architecture, functionality and attributes. We will investigate issues of trust in argumentation, in particular assumption doubt and differences between a notion of "good" (trustworthy) dependability cases and "convincing" (trusted) ones.