Myths and Barriers to the Introduction of Safety Cases in Space-Based Systems

Chris W. Johnson (1), Derek A. Robins (2)

(1) School of Computing Science, University of Glasgow, Glasgow, UK, G12 8RZ.
(2) SAIC, Houston, Texas, USA.

Abstract

Safety cases provide high-level support for the development of critical systems. They present an overview of the arguments and evidence that demonstrate a complex application is acceptably safe within a particular context of use. This approach offers particular benefits as organisations seek to procure services rather systems, for instance in government projects. Contractors can present safety arguments to explain the steps that have been taken to mitigate potential risks without providing low level proprietary information. Government agencies can inspect graphical overviews of safety arguments to identify potential weaknesses in the evidence that supports particular applications. There are, however, a number of concerns about the use of safety cases. For instance, the UK Hadden-Cave review has shown that safety cases can become 'tick-box exercises' if they are not supported by an appropriate safety culture. Other concerns stem from misunderstandings. For example, recent discussions about possible space based applications revealed that many engineers view safety argumentation as a means of reducing expenditure on existing forms of risk assessment and testing. This paper identifies the myths and rumours that jeopardise the development of safety cases to support complex space-based systems.

Introduction

Testing cannot guarantee the safe and successful operation of complex systems. There are clear ethical problems with evaluations that place operators or the public at increased risk. It can also be difficult to ensure that simulations and test-beds faithfully recreate key characteristics of an eventual working environment. Further problems relate to test coverage. It is infeasible to explicitly test every possible branch of execution embedded within many thousand lines of code. Similarly, the introduction of complex systems gradually influences existing operating practices through a myriad of subtle interactions. It can be difficult to recreate these adaptations in longitudinal tests.

Many of the same caveats can be applied to analytical techniques, including the use of formal or mathematical reasoning. It is often necessary to make strong assumptions about the environment in which a system will be used before model checkers or theorem provers can establish whether or not a design satisfies a particular set of requirements. Further problems arise because it is often unclear what those requirements should be until relatively late in the development cycle when, for instance, other systems have been integrated into a safety-critical application.

The limitations of product testing and abstract analysis have led many regulatory authorities to recommend process based techniques to support the engineering of complex applications. In particular, risk analysis has been used to identify and then mitigate the hazards associated with many safety-critical systems. By following approved techniques for risk assessment, developers cannot entirely eliminate potential flaws. However, these approaches help to guide and document the allocation of finite development resources. External auditors support these risk-based approaches by inspecting the artefacts that are generated during the application of a process based approach to systems development. Reliability centred maintenance techniques extend this approach into the everyday operation of complex systems; using hazard analysis to tailor maintenance intervals to the likelihood and consequence of component failures.

As might be expected, most organisations have integrated empirical testing into the risk-based development of safety-critical applications. Information gathered from empirical studies and from operation data can be used to inform hazard analysis. Lessons learned from previous adverse events can be used to identify new test cases for existing systems and proposed designs. The more recent techniques of resilience engineering can also be used to strengthen defences that have protected applications from previous hazards. This enables designers and developers to focus additional resources on those mitigations that have contributed most to safe and successful operation, rather than directing finite resources in response to bench tests that often do not replicate working environments.

The move from product based development to process-based approaches has also had a significant impact on regulatory bodies across many different countries (Kelly, 2008). From as early as the nineteenth century, governments had created specific legislative requirements for the operation of safety-related systems. These specified particular operating and performance characteristics that had to be met in order for an application to be approved. However, continuous innovation often meant that many of these requirements in the early 'Factory Acts' were obsolete, inappropriate and hard to apply. The development of process based regulation freed many statutory bodies from the need to maintain and enforce these laws. Rather than inspecting processes to determine whether or not they conformed to the prescribed characteristics, regulatory agencies began to focus on the documentation that was produced during the development processes that were intended to demonstrate that an application was acceptably safe.

However, this integration of product-based testing and process based maintenance and development has led to a proliferation of documentation across many safety-critical industries. This has reached the point where many organisations rely upon sophisticated document management systems to maintain and structure libraries of safety related information. The overheads associated with the creation and maintenance of these resources can be exacerbated as more and more systems rely upon the integration of sub-components from many different suppliers. Each of these different companies, typically, may use different processes and testing techniques to support the development of their sub-systems. This is a particular concern, as we shall see, for multi-national space related projects such as the International Space Station. Common concerns include the need to trace the implications of operational and test results across the design of complex systems. A failure of a single component can force the revision of risk assessments across many different applications that interact with that sub-system. It is non-trivial to ensure that this information is considered by all of the many different stakeholders in complex systems development.

The integration of diverse safety documentation from multiple suppliers has led many developers to consider using safety argumentation tools. These help to integrate different forms of analysis, including but not limited to Fault Trees, FMECA, HAZOPs, empirical and human factors testing, formal methods etc. Figure 1 shows these information sources as the evidence that can be cited by developers and operators in support of a claim that their system is acceptably safe. This diagram also shows how safety argumentation focuses resources on a goal-directed approach to development. Evidence is gathered to support the arguments in favour of particular goals or claims. In the context of this paper, a top level goal is assumed to be that the system is acceptably safe. However, many argumentation techniques also explicitly represent more detailed sub-goals or claims, such as 'software has been developed to an appropriate integrity level' or 'the probability of a high consequence hazard occurring is less than 1 in 10-6' etc.

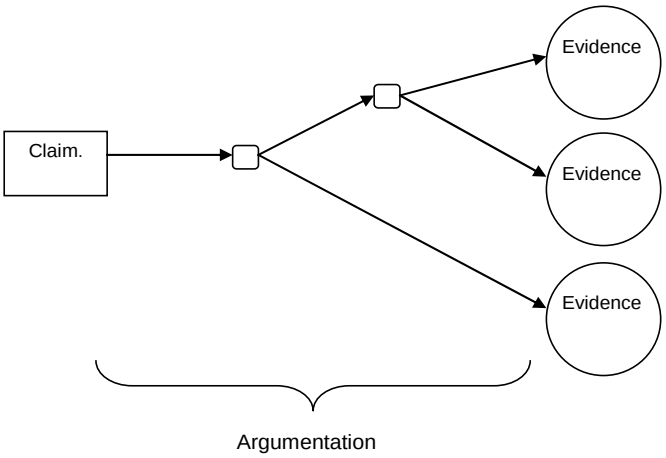

Figure 1: A High-Level View of Safety Arguments

Safety argumentation systems differ from conventional document management systems because the various sources of evidence contribute to specific arguments about the acceptability of a safety-critical application. This is important because if an error is found in a particular analytical technique or operational experience undermines the predictions derived from empirical tests then safety managers can immediately identify the implications for the overall argument that a system is safe to operate. In other words, they can trace the contribution that each item of evidence in Figure 1 makes to the overall argument about systems safety (Bloomfield and Bishop, 2010). It is important to stress that the approach sketched in Figure 1 is not simply intended to provide a map for safety arguments. These graphical structures are also intended to be the focus for

discussion between different stakeholders who can have very different viewpoints on the validity of particular safety arguments. It is also important to note that some safety case techniques do not rely on graphical overviews but instead exploit text based alternatives (Holloway, 2008).

UK Military Standard 00-56 reinforces the generic view illustrated in Figure 1 when it defines a safety case to be "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 environment". More recently, the US Presidential report into the Deepwater Horizon accident advocated the development of safety argumentation across the oil and gas industry (US Presidential Commission, 2011). The same approach illustrated in Figure 1 has also been applied to support security concern, for example ISO 15026 introduces the concept of a security assurance case that links evidence to arguments that a system meets particular security requirements. This has led to the promotion of dependability arguments as a generalisation beyond safety to consider the wider requirements and tensions in systems reliability.

It is important to stress that the use of safety argumentation techniques does not guarantee that an eventual system will be acceptably safe (Greenwell et al, 2006). However, the apparent success of this approach across many different industries has encouraged a number of national and international space agencies to consider the use of safety argumentation. The remainder of this paper describes a number of the myths that have been identified when talking to systems safety engineers about this application of the approach. Subsequent sections go on to distinguish these unsubstantiated concerns from the more justified barriers to the introduction of safety argumentation techniques.

An Introduction to Dependability and Safety Cases

The opening sections of this paper have introduced the background to safety argumentation techniques as a means of structuring and controlling the documentation that drives a process-based and goal directed approach to systems safety. Several alternate methods have been developed to support the application of safety argumentation. These include the Goal Structuring Notation (GSN) (Kelly and Weaver, 2004) and its derivatives (EUROCONTROL, 2006), as well as the Conclusions/Claims, Argument and Evidence approaches (CAE) (Johnson, 1999). Sophisticated tool support is also available, through the Adelard Safety Case Environment (ASCE) (Bloomfield and Bishop, 2010) and Artisan's GSN modeller. These techniques extend figure 1 in a number of ways.

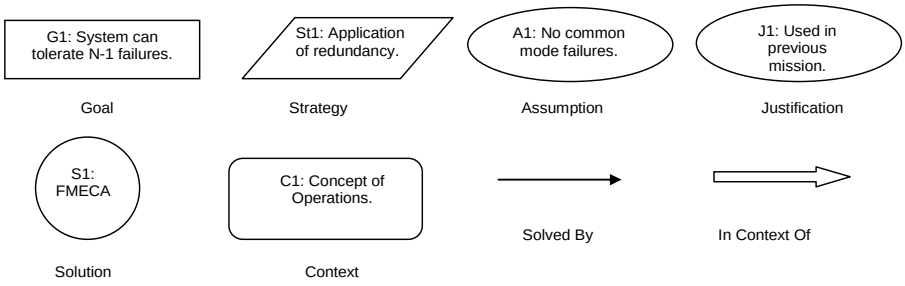

Figure 2: Components of Safety Argumentation Techniques

For example, Figure 2 illustrates a number of common syntactic components from the approaches cited above. As mentioned previously, a goal or claim represents an assertion that can be assessed as either true or false. For instance, a developer might assert that a system is 'acceptably safe during an emergency shut-down'. Although it might not be possible to derive conclusive proof of this goal, a regulator can either accept or reject the assertion. A strategy can be used to describe generic approaches to the arguments that are used in support of a goal or claim. For instance, reference to appropriate standards can be used in many parts of a safety argument. Similarly, many safety arguments exploit the strategy of combining a safety analysis of systems components with integration testing. A solution can be used to present the evidence that supports a goal or strategy. This is important because it provides a link between the high level argument structure embedded within GSN and CAE diagrams and the more detailed documentation provided by specific development techniques such as Fault Trees, FMECA, Formal methods etc. Regulators can inspect the evidence that supports particular solutions. If they are dissatisfied then it is relatively easy to trace the links from a solution to the higher level safety arguments that it supports. Figure 2 also presents a context node. This refers to the environment in which a system may eventually be deployed. Previous sections noted that safety arguments may not hold when an application is moved from one context to another. Assumptions are used to document areas of a safety

argument that are still to be supported by the evidence from particular solutions. It is important to explicitly identify these components because they indicate areas for further analysis or elements of an argument that depend upon evidence supplied by other stakeholders. Justifications help to document the reasons why a particular strategy or solution is appropriate. This is important when regulators or auditors may challenge the choice of particular development tools and techniques.

Figure 3 illustrates an application of the GSN approach to provide a high-level sketch of safety arguments relating to Extra-Vehicular Activities on the International Space Station (Johnson and Fodroci, 2011). As can be seen, the top level goal is to ensure that an ISS EVA is acceptably safe. This is decomposed into sub-goals. For instance, G2 states that all identified hazards have been eliminated or mitigated. This goal is influenced by a range of associated guidance documents that establish risk assessment as a primary means of ensuring the safety of NASA's EVA operations including NPD 8700.1 and NPR 8705.5. These documents appear in Figure 3 as part of the context for the GSA.

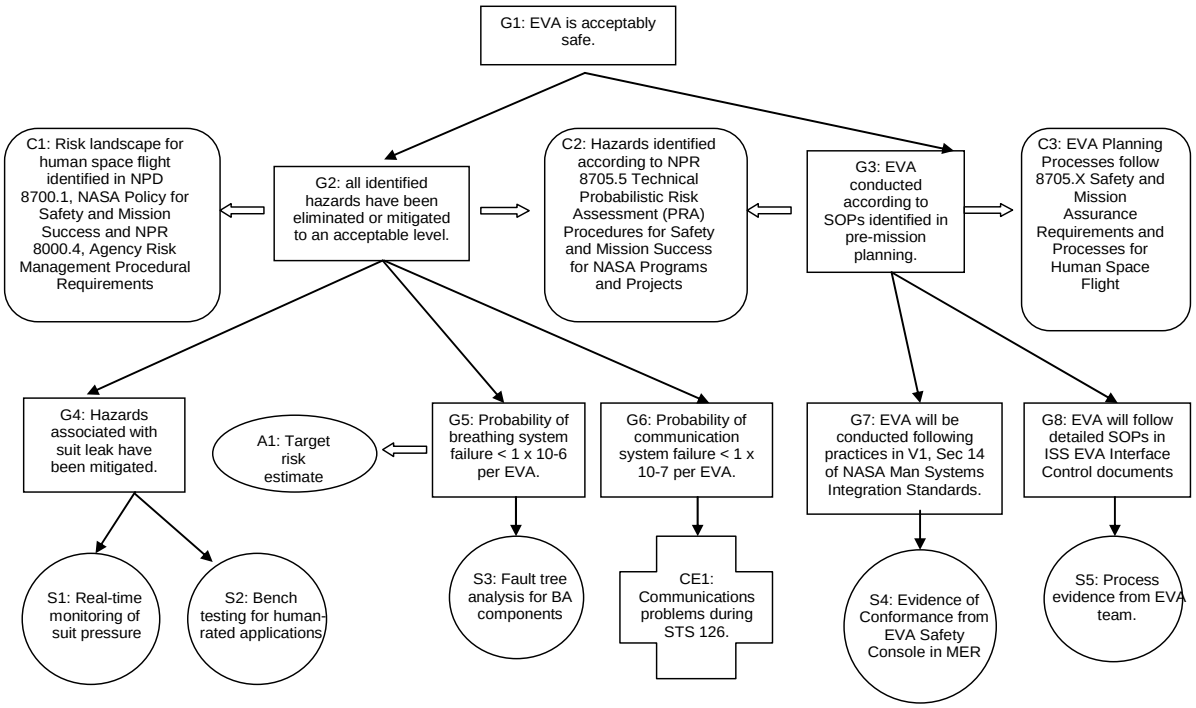

Figure 3: Initial GSN for ISS EVAs

The sub-goal G2 leads to a number of further claims. G4 asserts that the hazards associated with suit depressurizations have been mitigated. G5 states that the probability of a breathing system failure in less than 1 x 10-6 per EVA. G6 states that the probability of a communications systems failure is less than 1x10-7 per EVA. These reliability figures are estimates and hence an assumption A1 is used to show that further work is required to validate these goals. However, the GSN also shows the evidence that might be derived to support reliability estimates. In the case of G5, fault tree analysis might be used to compute failure probabilities for the breathing system. In the case of the communications system we have extended the GSN approach by introducing counter evidence (CE1). In our example, this refers to problems experienced during an EVA false alarm when an astronaut had his headset volume control knob inadvertently turned down. In this case, we might expand the GSN to consider human factors concerns in addition to the probability of failure from technical malfunctions. Figure 3 also shows how GSN diagrams can be used to integrate evidence from bench testing (S2) to support G4; supporting claims that risks of suit depressurisation have been mitigated. Real-time monitoring by the EVA Safety Console team in the Mission Evaluation Room (MER) can also help to mitigate this risk (S1).

The right-hand side of Figure 3 deals with operational planning and practice for EVA operations associated with the ISS. As with the other safety goals, this claim (G3) is influenced by the organisational context in which these activities are being conducted. In particular, NASA 8705.x guidance on Safety and Mission Assurance Requirements and Processes for Human Space Flight provides a framework for the safety argument (C3). More detailed sub-goals G7 and G8 make claims that EVAs will follow the practices outlines in Volume 1, Section 14 of the NASA Man Systems Integration Standards and that EVA Standard operating Procedures will follow the ISS EVA Interface Control Documentation. These arguments can be supported by evidence derived both from

the documentation and records of the EVA Safety Console and also from the logs and other documentation provided for EVA crew activities.

It is important to stress that Figure 3 only provides a partial sketch of the safety arguments that support EVA operations associated with the ISS. The specific documents identified within the context, such as NPD8700.1, might be superceded or revised. In which case, safety managers must reconsider the links between the revised documentation and the arguments made in a GSN. Similarly, there are several sub-goals that might be added – for example in terms of the interactions between design and operations or between the ground teams and the crew that help to mitigate any residual risks. The key point is that these diagrams act as a focus for discussion about the higher level safety arguments supporting complex systems. They provide a framework that integrates many different forms of evidence. They also help to focus attention on those areas of a safety argument that can be affected or undermined by contradictory evidence, such as the experience from STS126 illustrated in Figure 3.

Medium Term Influences on Space Industry

The changing financial context of space operations helps to explain the recent increased interest in the application of safety cases to multi-national projects. Many governments face significant fiscal pressures that have knock-on effects for their civil space agencies. This has led to the cancellation or curtailment of long term programmes, including NASA's Constellation initiative. ESA's spending in 2010 and 2011 has been frozen at approximately €3.7bn. Some member states, including Ireland, Portugal and Spain, have experienced considerable difficulties in securing their individual subscriptions. In the United States, the budget deficit has fuelled Republican hostility to the administration's plans for the integration between Federal and commercial space programmes. It seems likely that elements in Congress will try to cut subsidies for commercial human space flight from $6 billion over six years to $3 billion. These financial pressures arguably increase the importance of safety management systems. It is necessary to ensure that a lack of resources does not increase the hazards that have to be faced in many space operations. The graphical argumentation techniques presented in this paper provide powerful tools that help to ensure financial constraints do not have an adverse impact upon safety. For instance, the nodes in these diagrams can be used to trace the detailed impact of any decision to cancel the empirical testing of systems components. If the evaluations mentioned in S2 of Figure 3 came under threat then it is possible to reconstruct the impact that this would have at various levels of the safety argument from G4 to G2 and G1. Alternatively, attention might focus on the costs of meeting reliability targets expressed in G5 and G6; any savings here might arguably have less of an impact given that the reliability estimates are derived from the assumption A1. The point here is not to devise a detailed plan of cost reduction for Extra Vehicular Activities but to show how the same diagrams that can be used to structure safety documentation can also be used as a focus for discussion about the impact of cost reductions on the overall safety of space missions.

The financial pressures identified in the previous paragraph have combined to increase the importance of multi-national projects. Most notably, the end of the US Shuttle program has increased the reliance of the ISS on both Soyuz and ATV missions. This creates situations where safety depends upon the integration of sub-systems built using very different methods across many different partner nations. This can be illustrated by the simultaneous failure of all six Russian ISS central and terminal computers during STS-117. The loss of computational support affected the Russian Elektron Oxygen generator. Ground teams worked to provide an alternative back-up for the Elektron system. A plan was quickly developed and validated to install a hydrogen vent valve during an additional EVA. This enabled a new U.S oxygen generator to be brought on-line. However, the computational failures also affected attitude control for the ISS. Control moment gyros (CMGs) could be spun to counteract induced momentum during normal operations. However, these were insufficient to control the forces created by disturbances such as an orbiter undocking. In such cases, the CMGs must be taken offline and the station allowed to enter free drift. Once the Orbiter has undocked, Russian computer-controlled thrusters can be fired until control is returned to the CMGs. Without the Russian software for the on-board thrusters, the ISS relied on attitude control from the Orbiter's thrusters. This created a catch-22 situation where the ISS relied on the Orbiter to counteract any momentum imparted when that Orbiter undocked. If the Orbiter could undock then it was likely that the gyroscopes would quickly have become saturated and the only apparent way to avoid a loss of control would have been to fire the ISS thrusters which, in turn, depended on the failed computer systems. The Orbiter has been scheduled to return one week after the initial systems failure. Ground teams began to focus on using thrust from a Soyuz or Progress cargo ship after the departure of STS-117. The development and safety assessment of these plans illustrates the need to integrate technical input from the space programmes of multiple nations. As we have seen, safety argumentation techniques were deliberately developed to support the integration of diverse forms of safety analysis and their associated documentation. This is particularly important when, for instance, one nation makes extensive use of 'proven in use' arguments

while other agencies rely on the procedural approaches to safety that were introduced in the opening sections of this paper.

A further consequence of financial stringency has been an increasing interest in commercial support to assist Federal space programmes. This is particularly apparent in the Obama administration's recent support for the Commercial Crew Development (CDC) programme driven by organisation including the Commercial Spaceflight Federation**.** The President's 2012 NASA budget request to Congress includes $850m to help fund the initial development of several 'private' space vehicles. One justification for this initiative is the need to reduce US reliance on the Russian Soyuz vehicle for crew transfer to the ISS. A further argument in favour of this mixed approach is that NASA would then be free to focus its longer-term efforts on heavy lift Space Launch Systems (SLS) capable of taking crews beyond the ISS. Recent changes in the budget request have included increases in funding for the CDC but the amounts provisionally allocated to the SLS have remained the same. This integration of commercial and Federal space programmes mirrors developments in Russia and the European Space Agency as both seek partners to provide technical input and defray the costs of future missions. The key point here is that it is no longer possible to assume that commercial and state-sponsored agencies from a range of different nations will all use the same development processes. In some situations, there are Intellectual Property concerns over the disclosure of detailed engineering information. It is, therefore, critical that we find some common structure for those aspects of space systems engineering that have implications for mission safety when commercial platforms support multinational infrastructures, including the ISS.

Myths and the Introduction of Safety Cases

Given the widespread use of safety cases across several European industries, there have been several attempts to provide detailed evidence about their potential strengths and weaknesses for complex, multi-national space projects. For instance, NASA's Independent Validation and Verification Research Program has sponsored work to assess the utility of safety cases for the Ares abort fault detection, notification & response systems. As might be expected, the focus of this project was on the application of argumentation techniques to address the problems that arise in assessing the safety of software intensive systems. The Constellation program had previously decided not to adopt the approach because it was viewed as having a relatively low level of technology readiness. The IV&V research initiative was also intended to provide NASA with internal experience in the application of an approach that has not yet gained momentum in the US. It remains to be seen whether the Safety Case guidance that is being developed by this programme will support future NASA projects following the end of Constellation in the Authorisation Act of 2010.

The wider aims of the IV&V programme in addressing the perceived 'low technology readiness' of safety cases for space applications can be contrasted with the narrower technical objectives of NASA projects NCC2-1426 and NNA07BB97C (Basir, Denney and Fischer, 2009). One component of these initiatives has been to identify ways of automatically generating safety cases from the formal, mathematical proofs that support the development of safety-critical software. However, there is a meta-level safety argument that also supports the application of the methodology – providing evidence and assurance that the generation of safety cases is acceptable. As in the opening sections of this paper, the approach is justified in terms of the problems that arise in the application of product and process based methods for the certification of safety-software. The authors argue that existing standards, including DO-1798b "are typically process-oriented and require that code generators are qualified for application, often using an elaborate testing regime" (Basir et al, 2009). This is time-consuming and expensive. In contrast, formal reasoning techniques provide a 'product-based' approach by focussing on mathematical properties of a specification that are then used as evidence in the lower levels of a more general safety case.

A third example of the previous application of safety cases within space based engineering is provided by a project to create a software safety risk for legacy applications (Hill, 2007). The intention is to identify potential risks that might arise when existing code is re-used in future systems. The taxonomy was modelled on the SEI's taxonomy of risk factors (Carr et al, 1993). However, a key component of this work was the development of these potential hazards as a means of informing the development of future safety cases across a range of space-related software applications. This helps to ensure that problems with previous systems are not replicated in future applications. In this context, safety cases were promoted as a tool to support the application of the NASA Software Safety Standard (2004). In particular, the intention is to provide a method for developing the informal arguments that must demonstrate the acceptability of any legacy safety-critical computer system which is reused to monitor and control new hardware.

As can be seen, existing examples of safety cases within space-based systems have focused on support for the development and re-use of software applications. This is explained by the inherent difficulties that arise in the certification of these systems. Product testing cannot be used when there are many millions of execution paths through complex code. Process based development techniques also raise a host of questions in ensuring that code is tailored to particular operating environments. As we have seen, safety cases provide means to unite these approaches within common frameworks of argumentation. This focus on software certification has, arguably, had a negative effect in terms of wider perceptions. Not only is there a reluctance to exploit methods that have been developed in non-space related systems in Europe; there are further concerns about the wider utility of engineering techniques that are associated with the 'unique problems' that are in safety-related software engineering. The remaining paragraphs build on this analysis and present a number of further 'myths and objections' to the use of safety cases to support the development of space related systems.

It is important to stress that each of these initiatives must be set against a background of many hundreds of similar research projects commissioned by space agencies in Europe and North America. As can be seen from the preceding discussion, however, the teams involved are very clear about their mandate and the scope of their work. In spite of this, a recent initiative to expand understanding of safety cases within engineering teams at one NASA centre helped to identify a number of common myths about the potential application of this approach. None of these is unreasonable or irrational. They represent well justified concerns about the future use of the approach. They also arguably represent an area for improved communication between the proponents of these techniques and the front line, space engineers who might use them. However, these myths and misconceptions are important at a deeper level because they reflect the concerns and worries of space engineers at a time of profound change across the industry.

Mandatory Imposition. The first myth that engineering teams raised during the introduction to safety cases was that their use would become obligatory in safety-related space based missions. Previous paragraphs have explained the general perception that for many US missions the techniques were not at an appropriate level of technical maturity. Hence, there was no basis in fact that might substantiate this concern. The concern that safety cases might be imposed as a project requirement is interesting at many levels. It can be argued that it reflects the need to find new techniques to address the organisational and technical challenges that must be addressed in the medium term by many space agencies. Engineers are well placed to understand the strengths and weaknesses of existing tools. Previous sections have identified how the rise of multinational and commercial ventures is creating new engineering challenges for safety-related systems. These are compounded by the increasing technical sophistication of many potential projects together with the presence of legacy infrastructures and applications that cannot be replaced for financial reasons. In such circumstances, it makes perfect sense for engineering teams to expect the rapid introduction of new tools and techniques.

It might also be argued that concerns over the mandatory imposition of safety case techniques reflects communications issues between engineering teams and other levels of their organisation. Even in more 'normal' circumstances, multi-national projects must take coordinated decisions to introduce common techniques across the many different groups that contribute to major space missions. Financial and organisational uncertainty often undermines the traditional ways in which information is exchanged both between and within agencies. In such circumstances, engineers seem prepared to believe that they might be required to use a new approach without being consulted before hand.

An Alternative to Risk Assessment. During initial discussions with the engineering teams, it became apparent that there were a number of miss-conceptions about the role and scope of safety cases, rather than their potential application across space organisations. In particular, several individuals had assumed that argumentation techniques, including GSN, might be used to replace conventional forms of risk assessment. As we have seen, this is not the case. Typically, the findings from techniques including FMECA, PRA, HAZOPS etc might be cited as evidence in the lower levels of a safety case. Figure 2 illustrates the manner in which higher level arguments about the identification and mitigation of hazards can appear at higher levels in the structure. In both cases, safety cases are used as a framework to support the application of these existing techniques.

A number of explanations can proposed for this misconception over the scope of safety cases. For example, it might be argued that graphical and textual overviews of safety arguments can help project management teams to identify critical areas of concern during the development of complex systems. This, in turn, could help them to prioritise resources, including a reduced reliance on risk assessments in areas of 'marginal concern'. This explanation relies upon considerable insight into the application of safety argumentation techniques. It is difficult to sustain, given that the engineering teams had not met these approaches before. Further explanations focus on the many problems that arise in the application of conventional risk assessment techniques to complex

safety-related, space missions. Many of these systems are software intensive with considerable scope for human intervention both by flight crews and ground support teams. Existing PRA techniques are notoriously bad in terms of the support that they offer for software and for human reliability. The suggestion that safety argumentation techniques might reduce dependency on conventional risk assessment techniques may, therefore, reflect underlying concerns about our continuing reliance on these approaches across many space missions.

A Means of Implementing Cost Reductions. Previous sections have explained how the fiscal crisis facing many governments has placed considerable constraints on space agencies around the globe. It is, therefore, hardly surprising that new techniques might be interpreted as means of controlling and reducing costs. This misconception creates particular concerns because safety budgets are often difficult to defend when other areas of a programme are being axed. For example, it can be argued that expenditure on safety analysis adds little to the functionality of a programme. It can also be argued that safety studies are based on overly pessimistic or conservative assumptions; leading to mitigations and defences against potential hazards that are unlikely to arise during a potential mission.

In contrast, previous sections have argued that safety cases provide a higher-level framework for existing safety studies. They can create additional costs because development teams have to maintain the argumentation structures in addition to the risk assessments, empirical studies, lab tests and formal analyses that might otherwise be conducted. Any savings that do arise are as a result of improved safety management functions, different stakeholders get a better view of the role that each individual study plays in an overall argument about the acceptability of a complex system. At the same time, criticisms have been made about the abuse of safety cases (Haddon-Cave, 2009). As the level of detail and the scope of an argumentation structure can expand over time, it can also add unnecessary costs if development teams lose control of their framework. This can lead to a situation in which staff are reluctant to make changes to a complex system not so much because of the engineering implications but because they are unsure what those changes would mean to the underlying safety argumentation structure.

A Means of Implementing Skill Reductions. The engineering teams identified further misconceptions during preliminary discussions about the application of safety cases to support the engineering of space based systems. One set of concerns focussed on the role that safety cases might play in deskilling the engineering work force. At first sight it is difficult to determine where this misconception might have originated. Figure 2 illustrates the manner in which GSN and similar approaches help to reinforce safety engineering techniques by integrating them within a common argumentation framework.

One explanation for the suggestion that safety cases might help to deskill the existing engineering work force is that the introduction of a new technique helps to focus existing concerns amongst engineering teams. In other words, individuals were already concerned that this might be an organisational response to increasing funding constraints. This again illustrates how initial responses to safety cases tell as much about the general challenges to safety engineering across the space industries as they do about the technical strengths and weaknesses of the proposed approach.

A further explanation of concerns about de-skilling is that graphical or textual representations of safety arguments help to explicitly document issues that would otherwise have remained implicit within the development of complex space-based systems. The very simplicity of techniques such as GSN may also give a false impression that safety cases could provide a ‘silver bullet’ for a host of engineering problems. However, the proponents of these approaches are clear that they cannot replace the skills and expertise of space engineers. In this interpretation, a key contribution of safety professionals is to maintain an overview of the ways in which different forms of evidence contribute to overall claims about the safety of a complex system. This argument is hard to sustain, given that the proponents of safety cases would argue that these frameworks support the skills of safety engineers who are still needed to construct and maintain these documents, understanding the strengths and weaknesses of each of the more detailed approaches that contribute the evidence embedded within an argument.

Barriers to Safety Cases

It is important to stress that these myths can be balanced against genuine misgivings about the introduction of safety cases that were also raised during initial discussions about their application within space systems. Several of the engineers were concerned about the lack of expertise in the approach, not only within their organisation, but more generally across North America and throughout the space industry. This created a situation in which a small number of ‘local experts’ who understood the approach would quickly become

gatekeepers or guardians for the safety argumentation. This could create significant problems if those individuals became a bottle neck for update requests or a single point of failure in the case of illness. Further questions were raised about issues of scale when multiple agencies were contributing different forms of evidence to a safety argument. How could this be managed and supported b independent verification and validation? Finally, there was a concern that safety argumentation techniques might lead to a 'cut and paste' approach to safety. It would be very tempting to create a generic template for the relationship between testing and risk assessments that could be used in many different contexts but which were inadequately supported by underlying studies.

These concerns could not be dismissed as myths or misconceptions. They mirror many of the criticisms that have recently been published by the UK Ministry of Defence and the Haddon-Cave (2009) review into the loss of RAF Nimrod MR2 Aircraft XV230 in Afghanistan in 2006. This report is significant because safety cases were introduced in the UK Ministry of Defence for many of the same reasons that have been used in the earlier sections of this paper to justify their use across the space industry. Safety argumentation techniques were perceived to offer important means of structuring safety arguments across multi-partner projects including significant commercial input in response to Government requirements. In this case, BAE Systems worked with the MOD Nimrod Integrated Project Team and QinetiQ to develop the underlying safety arguments. The accident occurred when a leak developed during midair refuelling; fuel accumulated in the bomb bay where it was ignited by either an electrical fault or hot air leaking from a heating pipe. The subsequent review argued that the accident stemmed from "incompetence, complacency and cynicism" at various stages in the life of the aircraft and that the development of the associated safety case was the best opportunity to identify the design shortcomings.

However, this opportunity was lost; "If the Nimrod Safety Case had been prepared with proper skill, care and attention, the catastrophic fire risk to the Nimrod MR2 fleet represented by the Cross-Feed/Supplementary Conditioning Packduct and Air-to-Air refuelling would have been spotted and XV230 would not have been lost". However, the four years that it took to develop the argumentation did not yield valuable insights. In contrast, the safety case was a lamentable document riddled with errors. One of the reasons for this was the assumption that the Nimrod aircraft was 'safe anyway' and that the safety case simply documented rather than challenged this perception. Among the criticisms of the safety argumentation was that:

- "the project planning was poor;
- the personnel involved were insufficiently trained and inexperienced;
- the general approach was flawed from the outset, the task was wrongly regarded as essentially a documentary exercise;
- there was no sensible priority given to the high risks;
- there was no continuity of personnel;
- there was little operator input;
- the project management was inadequate;
- there was insufficient guidance for staff;
- the man-hours estimate was inadequate;
- the task was inadequately resourced;
- there was disagreement, confusion, and dissent between those involved as to how to proceed; etc."

As the project fell behind significant deadlines, there was pressure to proceed even though significant safety evidence was not derived. Instead, other sources of evidence were used that were not adequate for the role that they played in the overall safety case. There were no proper internal reviews even though the IPT was told that these had taken place; "there was strong commercial motivation to finish the Nimrod Safety Case by the deadline at all costs".

Only an optimist would claim that these caveats could never be applied to the proposed use of safety cases as a means of addressing the diverse challenges of next generation space missions. However, this does not undermine the utility of the approach. The key lesson from the Hadden Cave review was that safety cases are not a panacea. In contrast, their success relies on the creation and maintenance of strong organisational cultures supported by appropriate Safety Management Systems. The irony is that these are precisely the attributes that can be undermined in times of profound organisational change or uncertainty. The myths and misconceptions identified by space engineering teams provide cogent insights into the impact that the rapid refocusing to space policy is having upon the human resources that underpin all future missions.

Conclusions

Safety cases provide high-level support for the development of critical systems. They present a graphical or textual overview of arguments and evidence, which can demonstrate that complex applications are acceptably safe within a particular context of use. This approach offers particular benefits as organisations seek to procure services rather systems, for instance in government projects. Contractors can present safety arguments to explain the steps that have been taken to mitigate potential risks. Government agencies can inspect graphical overviews of safety arguments to identify potential weaknesses in the evidence that supports particular applications. There are, however, a number of concerns about the use of safety cases. Some of these are based on previous experience. For instance, the UK Hadden Cave review has shown that safety cases become little more than 'tick-box exercises' when they are not supported by an appropriate safety culture.

Other concerns stem from misunderstandings. For example, recent discussions about the possible application of the approach in space based systems revealed that many engineers view safety argumentation as a means of reducing expenditure on existing forms of risk assessment and testing. This paper has identified some of the myths and rumours that can undermine the eventual application of safety cases to support the development of complex space-based systems. Most of these concerns reflect misunderstandings about the scope and technical support provided by safety argumentation techniques. The closing sections of the paper go on to argue that these myths also provide valuable insights into the concerns and uncertainty that arises at times of rapid change in complex, safety-related industries.

Acknowledgement

The work described in the paper has been supported by the UK Engineering and Physical Sciences Research Council grant EP/I004289/1.

References

N. Basir, E. Denney and B. Fischer, Deriving Safety Cases for the Formal Safety Certification of Automatically Generated Code, Proceedings of the First Workshop on Certification of Safety-Critical Software Controlled Systems, Electronic Notes in Theoretical Computer Science, 238(4), 28 September 2009, Pages 19-26,

RE Bloomfield, PG Bishop, Safety and Assurance Cases: Past, Present and Possible Future? In F. Redmill and T. Anderson (eds.), Making Systems Safer: Proceedings of 18th Safety Critical Systems Symposium (SSS'10), 51-67, Springer Verlag, London, 2010.

M.J. Carr, S.L. Konda, I. Monarch, F.C. Ulrich, C.F. Walker, Taxonomy-Based Risk Identification, Software Engineering Institute Technical Report, CMU/SEI-93-TR-6, Carnegie Mellon University, Pittsburgh, Pennsylvania, 1993.

EUROCONTROL, Safety Case Development Manual, Technical report DAP/SSH/091, Brussels, Belgium, 2006.

W.S. Greenwell, J.C. Knight, C.M. Holloway and J.J. Pease, A Taxonomy of Fallacies in System Safety Arguments. In Proceedings of the 2006 International System Safety Conference, International Systems Safety Society, 2006.

C. Haddon Cave, An Independent Review Into the Broader Issues Surrounding the Loss Of The RAF Nimrod MR2 Aircraft XV230 In Afghanistan IN 2006, Report HC 1025, Her Majesty's Stationery Office, London, 2009.

J. Hill, A Software Safety Risk Taxonomy for Use in Retrospective Safety Cases, In Proceedings of the 31st IEEE Software Engineering Workshop, 2007. Columbia, MD, Feb. 8 2007.

C.M. Holloway, Safety Case Notations: Alternatives for the Non-Graphically Inclined? In C.W. Johnson and P. Casely (eds.), Proceedings of the IET 3nd International Conference on System Safety, IET Press, Savoy Place, London, 2008.

C.W. Johnson, A First Step Towards the Integration of Accident Reports and Constructive Design Documents. In M. Felici, K. Kanoun and A. Pasquini (eds.), Computer Safety, Reliability and Security: Proceedings of 18th International Conference SAFECOMP'99, 286-296, Springer Verlag, 1999.

C.W. Johnson, A Handbook of Accident Investigation, Glasgow University press, Glasgow, 2003. Available on http://www.dcs.gla.ac.uk/~johnson/book, accessed February 2011.

C.W. Johnson and M.P. Fodroci Promoting Resilience in Human Space Flight at a Time of Fiscal Pressure, Submitted to the International Association fort he Advancement of Space Safety 2011, Paris, France, ESA/NASA. 2011.

T P Kelly, Can Process and Product-based Approaches to Software Safety be Reconciled? In F. Redmill and T. Anderson (eds.), Proceedings of 16th Safety Critical Systems Symposium (SSS'08), 3-12, Springer Verlag, London, 2008.

T P Kelly and R A Weaver, The Goal Structuring Notation - A Safety Argument Notation. In Proceedings of the Dependable Systems and Networks 2004 Workshop on Assurance Cases, July 2004.

NASA Office of Safety and Mission Assurance, NASASTD-8719.13B Software Safety Standard w/Change 1, 2004.

US Presidential Commission, Deep Water: The Gulf Oil Disaster and the Future of Offshore Drilling, Washington DC, USA, January 2001.
Available from: http://www.oilspillcommission.gov/sites/default/files/documents/DEEPWATER_ReporttothePresident_FINAL.pdf

Biography

Chris.W. Johnson, DPhil, MA, MSc, FBCS, CEng, CITP,
School of Computing Science, Univ. of Glasgow, Glasgow, G12 8RZ, Scotland, UK.
Tel +44(141)3306053, Fax +44(141)3304913, Johnson@dcs.gla.ac.uk, http://www.dcs.gla.ac.uk/~johnson

Chris Johnson is Professor of Computing Science at the University of Glasgow in Scotland. He heads a small research group devoted to improving the reporting and analysis of incidents and accidents across safety-critical domains ranging from healthcare, to the military to aviation and rail.

D.A. Robins,
SAIC, Houston, Texas, USA.
derek.a.robins@nasa.gov

Derek Robins is...