On Arguing The Safety Of Large Systems

Gordon R. Stone

Compucat Research Pty Limited
14 Wales Street, Belconnen 2617, Australian Capital Territory

gordon.stone@defence.gov.au

Abstract

The Occupational Health and Safety (Commonwealth Employees) Act (OH&S Act) and technical regulations in the Royal Australian Navy (RAN) both place obligations on the Department of Defence in relation to the safety of its systems. The RAN's technical regulation system requires that evidence of safety be provided for new acquisitions. A safety case is preferred for major systems.

The upgrade of a frigate belonging to a class first commissioned over 35 years ago provides significant challenges to the provision of a safety case that should meet the requirements of the OH&S Act and the RAN's technical regulations. This paper examines some of those challenges and discusses techniques and tools that are being applied to them.

Keywords: Safety case, safety argument, GSN.

1. Introduction

The OH&S Act requires that the Commonwealth provide a workplace that is without risk to the health and safety of Commonwealth Employees and third parties in or near the workplace. In addition the (Australian) Naval Technical Regulations require that safety evidence be provided for the products of each new acquisition project (Australian DoD, 2003), with a Safety Case the preferred form of evidence for major acquisitions. These two obligations are complementary, although they may result in different approaches to the Safety Case (Robinson 2003). A prudent safety manager will attempt to satisfy both the technical regulations and the common law requirements.

The Australian Department of Defence is upgrading its Adelaide Class guided missile frigates (FFGs). The Adelaide class is based on the US Navy's Oliver Hazard Perry class FFGs, the first of which was commissioned in 1977. Four RAN FFGs were built in the USA, the first commissioning in 1981. Two more of these ships were built in Australia.

The design of the Oliver Hazard Perry class was based largely on then fielded equipment. Therefore, the design of the ship class dates back to the 1970s and the design of some equipment dates back to the 1960s.

The safety information relating to some original equipment is limited, partly because safety programs were not as demanding when the equipment was built and partly because some of the original data were not retained over time. However, the ships have been and are operated under a workplace safety regime called NAVSAFE. NAVSAFE includes detection of hazards, temporary measures to avoid mishaps until the hazard is treated, and hazard treatment, including design changes where required.

The FFG upgrade includes some equipment that was designed in the 1990s and some software that completed design in the mid 2000s. In addition, modifications are made to existing equipment of earlier design.

Several vendors are providing the systems, equipment and software that are being introduced to the FFGs through the upgrade. The vendors are from several nations, including Australia. Some of the items new to the FFGs already exist and some are being developed specifically for the FFGs.

The project to upgrade the FFGs has a system safety program in place. This program requires the prime contractor to conduct a system safety program in accordance with a recognized US military safety process standard, MIL-STD 882C. There is also a contractual requirement for the contractor to enable the Commonwealth to meet its obligations under the OH&S Act.

2. Challenges

The following are challenges to development of a Safety Case for the upgrade to the FFGs:

1. The scarcity of safety data for many of the Configuration Items (CIs) in the existing ships.
2. Different acquisition strategies for different types of equipment.
3. The number of items that need to be considered.
4. The availability of Defence resources to review the safety evidence offered by the contractor.

The scarcity of safety data for many of the CIs in the existing ships is a significant challenge. Many of the companies that designed and manufactured the equipment no longer exist or now pursue different lines of business.

In many cases, the type of safety data now sought was not created when the equipment was originally developed. In most cases, safety data are now unavailable for equipment developed in the 1960s and 1970s. However, it is clearly established that the weapon systems were subjected to a stringent safety verification program when they were developed by the US Navy and the platform systems were built to US Navy standards.

Modifications to the ships since their commissioning have been made by inclusion of modification kits developed by the US Navy or development and inclusion of RAN modifications. The resultant safety risk depends on the system safety program of the organizations developing the modifications.

There is the possibility that unidentified hazards still exist due to the ship design. Long periods of operation without mishap may have occurred because the operational conditions required to expose one or more hazards have not previously been encountered. Operation of the ships under the NAVSAFE regime ensures that any mishaps will result in treatment of the hazard that caused them.

The different acquisition strategies for the new systems, equipment and software include purchase of existing items, either military or commercial; purchase of modifications to existing items and incorporation thereof; and new developments. The types of safety data for each of these acquisition strategies will generally be different.

The upgraded FFG comprises over 15,000 CIs. The project for the upgrade does not have sufficient resources to address all CIs within a reasonable time. Therefore, a risk-based approach is necessary. The resolution to which the ship systems need to be considered will be based on risk assessments. Due to resource constraints, these risk assessments will be less rigorous than formal safety arguments, although they must be sufficiently well founded to establish duty of care.

The Defence resources available for the FFG upgrade project are limited. In addition, specialized knowledge and experience is required to review safety issues relating to many of the systems. Therefore, an approach to the Safety Case that provides the equivalent of template arguments should enable a variety of personnel to review the contractor's safety data and reliably determine whether it substantiates claims about acceptable safety risk.

The following key issues were derived from the above challenges:

1. How to determine when the safety argument for the FFGs is complete to the extent required to ensure that the safety risk is acceptable.
2. The items that should be considered as needing an in-depth safety argument at the time that the upgrade is accepted.
3. How to ensure that all the safety evidence for each item that needs to be considered has been provided.
4. How to provide a concise summary of the safety status for such a large argument.

To determine that the safety argument for the FFGs is sufficiently complete to ensure that the safety risk is acceptable requires an argument framework that reflects the hierarchical breakdown of the ship into its component systems, equipment and software. Figure 1 illustrates the top levels of the ship system breakdown structure.

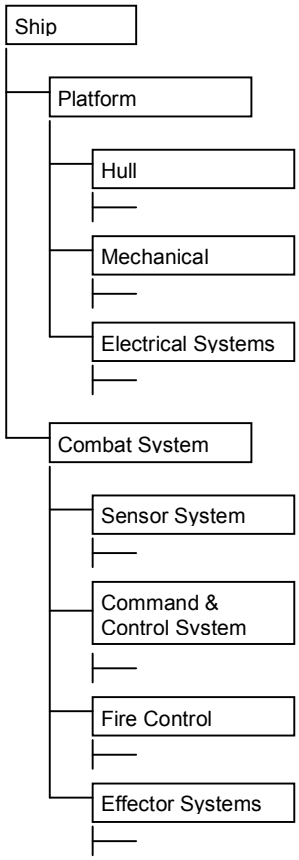

Figure 1: Example Ship System Breakdown

The argument structure must also enable consideration of functions that transcend individual systems, equipment and software. The completeness of the argument can be assessed by whether the argument has been developed to sufficient depth in each branch of the hierarchical breakdown structure to ensure that the safety risk of the remainder of the branch is well understood and acceptable.

The approach using hierarchically structured arguments and safety risk assessments as described above is the basis for deciding which items that should be considered as needing an in-depth safety argument at the time that the upgrade is accepted.

Determining that all of the safety evidence that should have been provided for each item has been provided is complicated by the need for domain specialists to review the evidence. One possible technique would use the system safety team to review the completeness of the evidence and the domain specialists to assess its adequacy.

For historical reasons, the safety evidence was not all developed and presented to the Commonwealth as the design of the upgrade progressed. Therefore, the system safety team needed to develop a concise expression of the types of evidence that the contractor should provide. The chosen approach was a set of generic arguments that can be instantiated as required for each CI. The instantiated arguments are used to test whether all the right types of information have been provided. Domain expertise is used to determine whether the evidence is adequate to support the argument. This approach is discussed in more detail below.

3. Approach

The approach to coping with the challenges and answering the key questions was to adopt a framework for evaluating the evidence delivered from the contractor's System Safety Program. The framework comprised the following elements:

1. Rules for selecting CIs for the safety argument for the upgrade.
2. A set of standard generic arguments that are deemed to be of suitable quality.
3. A means of instantiating an appropriate generic argument for each CI to be considered.
4. A means of assessing each instantiated argument, where the assessment is whether the argument holds or not.
5. A means of aggregating the assessment of individual instantiated arguments.

Rules for selecting CIs in the overall upgrade argument are as follows:

1. All CIs introduced to the FFGs by the upgrade are included.
2. All legacy CIs (items already in the FFGs) that were modified by the upgrade are included.
3. All legacy CIs that had their conditions of use changed by the upgrade are included.
4. All other CIs are excluded from the overall upgrade argument, but will be considered at another time and in another forum.

The term 'conditions of use' represents all aspects of installation, services, interfaces, operation and maintenance of the item. It also includes the legislative and regulatory framework relating to the safety of the item.

The concept of a set of generic arguments is based on safety case patterns (Kelly, T., McDermid, J., 1998), although the level of the arguments and the arguments themselves are not necessarily the same as discussed in the reference.

The generic arguments used in the upgrade are based on acquisition strategy for each CI, since the types of safety evidence that are likely to be available for each CI largely result from the acquisition strategy for the CI. The following acquisition strategies were used in the upgrade:

1. Legacy.
2. Modified.
3. Non Developmental Item (NDI).
4. Re Use Software Item (RUSI).
5. Developmental System Level Configuration Item (SLCI).
6. Developmental Hardware Configuration Item (HWCI).
7. Developmental Computer Software Configuration Item (CSCI).

There are different variants of some of these arguments. For example, the argument about the safety of legacy CIs has been divided into two as a matter of convenience. One argument is about legacy CIs that do not have changed conditions of use; the other is about legacy CIs that have changed conditions of use. Although the two could remain as a single argument, it was found that the ability to comprehend separate arguments was greater than a single argument.

There are also some alternatives within arguments. For example, a CI may have no safety critical or safety related functions. However, there may be safety issues related to its construction or installation, such as sharp edges or weak mountings. Its construction not only influences arguments about its safety as delivered and installed, but also arguments about maintaining it safely, maintaining it to be safe, and arguments about its disposal.

There are also generic arguments that can be used to support the arguments based on acquisition strategy. Some of these are:

1. Argument by appeal to a competent third party.
2. Argument by comparison.
3. Argument by licensed operator.
4. Argument by licensed repairer.

Consideration of both types of generic arguments above established that a relatively small set of such arguments is sufficient to address the entire upgrade.

The SLCIs, HWCIs and CSCIs in the upgrade are diverse in nature. They are also provided by several different vendors. Therefore, the generic arguments need to be tolerant of different ways that safety may be argued in detail. For example, the actual technique used to identify hazards is not crucial as long as it is credible, defensible, and appropriate to the problem to which it is applied; and personnel participating in the hazard identification have suitable domain competencies and suitable competencies in applying the technique. Therefore, the types of evidence can be prescribed in the generic argument. The merits of the evidence must be assessed by one or more domain specialists.

For historical reasons, the generic arguments were developed as the upgrade was nearing the completion of its development. This did not allow many of the benefits of developing phased safety cases to be realized. However, the arguments address the development and use of the CIs, and provide a framework for arguing safe disposal.

The acceptability of the generic safety arguments will be established by peer review and review by the technical regulators.

Consideration of the way that the generic arguments may be instantiated for the various types of CIs and acquisition strategies established that there will be a relatively small set of instantiated arguments required to support the entire argument for the safety of the ship. Therefore, once the generic arguments are complete, they can be instantiated to form the required set. Currently generic arguments will be instantiated by copying and editing.

For the argument for any CI to hold good, it is necessary to demonstrate that the argument itself is sound and that suitable evidence is provided to support the argument. To demonstrate that the argument is sound, it is necessary to demonstrate that the logic of the generic argument is correct, an *a priori* activity, and that it has been instantiated correctly for the CI. To demonstrate that the evidence supports the argument, it is necessary to demonstrate that the evidence is complete, that is each item of evidence required by the argument has been reviewed, and that each item of evidence supports the goal that it purports to support.

The merit of each item of evidence will be assessed by one or more persons competent in the domain(s) in which the evidence exists. In the parlance of the RAN's technical regulatory framework, Design Acceptance Representatives or their delegates carry out the assessment. A repository for the assessments is required. The repository needs to identify the evidence, identify the assessor, state the date that the assessment was made and state the results of the assessment.

Where an instantiated argument applies to more than one CI, the argument for each CI need only contain a reference to the instantiated argument and the assessment. It is not necessary to copy the instantiated argument for every CI that uses it.

The assessments of low-level arguments need to be aggregated into higher-level arguments so that the merits of the entire argument can be assessed. If the argument for the entire FFG is not substantiated, it is a useful to be able to determine the evidence that is deficient or not provided.

4. Tools Selection

When work commenced on developing the generic arguments, it was clear that, to achieve a consolidated safety argument for the ship, a highly structured argument technique was required. Goal Structuring Notation (GSN) was chosen to represent the argument. Of particular importance was the ability to represent the elements of the arguments and the relationship among the elements (Chinneck et al, 2004). This is a strong feature of GSN. Arguing the safety of each SLCI, including the ship, can be achieved by arguing the safety of CIs comprising that SLCI and the interactions among them.

The project team evaluated various tools available at the time for the development of arguments in GSN. Of particular concern was the ability of the tool to support:

1. hierarchical decomposition of arguments into manageable sized fragments,
2. compact representation of arguments,
3. extensive editing, and
4. electronic distribution for review and comment.

At the time, the number of tools and their sophistication did not meet all of the above requirements, in particular, the ability to electronically distribute arguments for review and comment. Electronic distribution using the then available tools was unattractive because reviewers would have needed another tool, additional licences would have been required and the tools would have had to be installed on different Defence networks. In addition to electronic distribution, compact representation was a significant issue. GSN is very easy to follow, but it is relatively sparse. It was considered that a more compact representation that preserved the fundamental structures of GSN could be used to develop and review the arguments, and then import them into a 'proper' GSN tool when they are stable.

The approach adopted was to build a set of tools based on Excel spreadsheets. Excel was chosen because it is one of the standard Defence office tools and all of the potential argument reviewers can use it reasonably well.

5. Tools

The concept underlying the tools is a worksheet that contains the arguments and a supporting worksheet that contains the attributes of the objects in the argument. An additional merit worksheet is provided to capture and aggregate assessments of Solutions.

Figure 2 illustrates the argument worksheet. The argument worksheet presents the argument in an indented list format, where the depth of indenting and the types of objects together show the structure of the argument.

Goal	X is safe			
	Strategy	Argue over Y and Z		
	START	Goal	Y has an acceptable safety risk	
			Solution	Evidence
			Solution	Evidence
	AND	Goal	Z has an acceptable safety risk	
			Strategy	Argue over ….

Figure 2: Argument Worksheet

The objects in the argument spreadsheet are equivalent to GSN objects. The type of each object is written alongside it. The object types are colour coded to facilitate recognition.

Guidewords are written alongside Goals to indicate how they are logically combined. These guidewords are derived from the text of the Strategy logically preceding them. The word 'cohesive' results in the logical connective 'AND', 'diverse' results in 'OR' and 'choose' results in 'ALT'. The guideword 'START' is written next to the first Goal in each logical combination. The guideword 'ALT' enables generic arguments to contain alternatives that are selected when the argument is instantiated.

Other guidewords are derived from the text of the Goals. These guidewords are 'AWAY', to indicate an away goal, and 'RECUR' to indicate that there will be more than one instance of the subject of the Goal, and each instance will have its own Goal when the argument is instantiated. Away goals are identified by the text 'Away Goal' at the beginning of the Goal's text. Recurrent goals are identified by the words 'Rule: One Goal per' anywhere in the text of the Goal.

Objects on the argument worksheet sheet are hyperlinked to the first cell in the same row in the attributes sheet. The hyperlink is used to quickly access the attributes.

The attribute sheet contains columns, one column per attribute. Each attribute is equivalent to a GSN attribute. Text wrapping is used in each column. The width of the columns depends on the number of attributes.

The merit worksheet is created automatically using the structure and objects copied from the argument worksheet. It should be created when all of the Solutions have been addressed. Figure 3 illustrates the main elements of the merit worksheet that are used for assessment. Assessments are colour coded to assist with recognition.

Obj Number	Obj Type	Logic Conn	Obj Assess	Obj Text
1	Goal		N	G1
1.1	Strategy		NA	S1
1.1.1	Goal	START	N	G1.1
1.1.1.1	Solution		N	S1.1
1.1.2	Away Goal	AND	Y	Away Goal: 1.2
1.13	Goal	AND	U	G1.3
1.13.1	Strategy		NA	S1.3.1
1.13.1.1	Goal		U	G1.3.1.1
1.13.1.1.1	Solution		U	S1.3.1.1.1

Figure 3: Merit Worksheet

Additional fields not shown in Figure 3 are the assessor's name and the date of the assessment.

The tools consist of a set of macros that provides on request automated:

1. highlighting,
2. addition of guidewords,
3. hyperlinking,
4. addition of an autofilter,
5. sizing of columns, and
6. font size selection.

In addition, the set of macros provides the ability to, on request:

1. hide and unhide rows,
2. parse the argument structure,
3. aggregate assessments of Solutions,
4. format data derived from a proprietary GSN tool, and
5. format data for input to a proprietary GSN tool.

Hiding and unhiding rows allows the user to focus on selected parts of the argument. The hide rows capability hides all rows between the selected Goal and the next Goal at the same or higher level. The beginning of the row containing the Goal used for selection is coloured grey to indicate that there are hidden rows under the Goal. There is no limit to the number of Goals that can have hidden rows. There is also no restriction on hiding Goals that have hidden rows.

Parsing the argument structure requires that each Goal that is not an away goal or is not immediately satisfied by one or more Solutions has a Strategy, and each Strategy contains a keyword that allows the logical connectives to be determined. It also requires that Goals that are away goals and/or recurrent goals be identified by specific words in the Goal text. With these restrictions in place, the parser can detect structural defects in the argument. The above restrictions are not required by GSN.

Aggregating the assessments of each instantiated argument enables a summary of the validity of the argument to be made. Aggregation is performed by aggregating the assessments of the Solutions to make an assessment of their parent Goals, then aggregating the assessments of Goals to make an assessment of their parent Goals.

The assessment of each Solution can be one of U (Unknown), N (Not satisfactory) or Y (Yes, it's satisfactory). All assessments for Solutions are given the default value of null. The algorithm to aggregate the assessments of Solutions is as follows:

1. Null for any Solution causes the parent Goal to have an assessment of N.
2. N for any Solution causes the parent Goal to have an assessment of N.
3. U for any Solution will not overwrite an N for the parent Goal, but will otherwise cause the parent Goal to have an assessment of U.
4. Y for any Solution will not overwrite an N or a U for the parent Goal, but will otherwise cause the parent Goal to have an assessment of Y.

The algorithm to aggregate the assessments of Goals depends on whether they are logically combined with AND or OR.

The algorithm for Goals combined with AND is as follows:

1. N for any Goal causes the parent Goal to have an assessment of N.
2. U for any Goal will not overwrite an N for the assessment of the parent Goal, but will otherwise cause the parent Goal to have an assessment of U.
3. Y for any Goal will not overwrite an N or a U for the assessment of the parent Goal, but will otherwise cause the parent Goal to have an assessment of Y.

The algorithm for Goals combined with OR is as follows:

1. Y for any Goal causes the parent Goal to have an assessment of Y.
2. N for any Goal will not overwrite a Y for the assessment of the parent Goal, but will otherwise cause the parent Goal to have an assessment of N.
3. U for any Goal will not overwrite a Y or an N for the assessment of the parent Goal, but will otherwise cause the parent Goal to have an assessment of U.

Formatting data from/for the proprietary tool allows arguments to be brought into the spreadsheet tools where they can be developed and reviewed rapidly, while preserving the ability to draw GSN when the arguments are stable. The proprietary tool, among other things, draws GSN diagrams from a hierarchically structured database. Although it provides many of the facilities required by the FFG upgrade project, it proved to have some limitations in terms of the criteria discussed in Section 4. Since a significant amount of work had been put into developing arguments in the proprietary tool, and since it is intended to put the arguments back into the proprietary tool when they are stabilized, macros to format data output from the proprietary tool and format data for input to the proprietary tool were required.

6. Configuration Management of Arguments

Configuration management is required during development of arguments and once arguments have achieved approved or agreed status.

One of the strengths of GSN is the ability to modify arguments to meet changing circumstances. A well constructed argument allows a competent person to readily identify where changes are necessary and make them. Modifying arguments is most effective when the arguments are well constructed.

For generic arguments, such as those being developed for the FFG upgrade, it is important that all logical paths are developed, for logical paths that are unnecessary in the current circumstances may become necessary in future circumstances. If such paths are not present in the generic arguments when modifications are required, different personnel may arrive at different modifications to the equivalent instantiated arguments, increasing the risk of error. However, consideration of currently unnecessary logical paths can substantially increase the complexity of arguments to be used in current circumstances, with consequential increased risk of error by the developers and reviewers.

Proprietary tools may provide robust configuration management capabilities when developing and maintaining arguments. For example, a proprietary GSN tool based on DOORS© should have sound configuration management capabilities through the inherent DOORS© functionality. Unfortunately, the tools developed for the FFG upgrade do not have inherent configuration management functionality. Configuration management is the responsibility of each person managing each argument.

7. Limitations

The following are limitations of the spreadsheet tools:

1. The set of macros does not draw GSN diagrams.
2. Configuration management of the arguments is not assured.
3. Arguments developed using the spreadsheets are generally too verbose for representation in GSN.
4. The development of attributes in conjunction with arguments is cumbersome.

The primary reason for using GSN is the clarity of arguments expressed in GSN. The clarity is reduced in the spreadsheet version. However, clarity can be recovered by importing the data into a tool that draws GSN diagrams. The capability to exchange data with a GSN tool has been developed.

The robustness of configuration management in the spreadsheet environment depends on the goodwill and diligence of the personnel developing and maintaining the arguments. Applications that draw GSN diagrams and have sound configuration management capabilities exist, and these are more suitable repositories for the final arguments. Unfortunately, the configuration management data used whilst developing the arguments in the spreadsheet tolls may not be able to be imported into proprietary tools with the GSN data.

The spreadsheet format does not impose the same restrictions on the quantity of text used for each object. This can be beneficial when arguments are under development. However, translation to diagrammatic form may require substantial text pruning. Once arguments are constructed and reviewed, pruning the text may invalidate the agreed status of the argument. There is no easy solution at present.

Development of attributes in conjunction with arguments is cumbersome, since editing an argument generally changes the row numbers of its objects. This causes the hyperlinks to have incorrect references. Although existing

hyperlinks are destroyed each time that hyperlinking is carried out, the process to align attributes and arguments can be tedious. The problem can be ameliorated by developing the argument to a point where it is stable and then adding the attributes. It can also be ameliorated by using a column containing the original row numbers on both the argument and attribute worksheets.

8. Practical Difficulties

The greatest difficulties experienced to date are as follows:

1. Finding time to develop the generic arguments. This is not an artefact of the tools or approach; developing sound arguments is time consuming. In some ways it is a vindication of the approach, since it obviates the need for reviewers of contractor safety data to make similar efforts when assessing contractor safety data.
2. Finding personnel with the right mindset to develop arguments. GSN is a very effective tool for reviewing arguments; however, development of GSN arguments is not an intuitive activity. Generally, people developing GSN arguments required a significant amount of time familiarising themselves with the technique before they became effective.
3. Organising peer review of arguments. Resource constraints within Defence limit the availability of appropriate resources for review of the arguments. Peer reviewers require some training in GSN. They also require time to give careful consideration to the arguments. There are very few people in Defence with appropriate domain knowledge that meet these criteria.
4. Managing the configuration of the arguments during development. There is no real solution, apart from personal diligence, using the spreadsheet tools. Proprietary tools provide better configuration management capabilities.

9. Conclusion

Spreadsheet representation of arguments for assessing the merits of safety evidence relating to very large systems has provided some benefits. A final assessment of the success of the approach cannot yet be made, since application of the technique to large quantities of contractor data has not yet occurred for programmatic reasons.

A substantial benefit of the aggregation of safety arguments as outlined above is that it will present a concise view of the where system safety is deficient or insufficiently well known. Reviewers can navigate from high level nodes worthy of attention to determine what makes them worthy of attention. This can be useful for assessing overall safety risk and planning remedial activities.

The application of GSN arguments to systems that have been developed or are nearing the completion of development may not provide a great improvement in assessment of system safety. Current indications are that, unless the safety evidence required to support the arguments is identified at the commencement of the development program, it may not be collected during development. That is, there may be a mismatch between the evidence required to support the arguments and the evidence collected, and arguments may fail as a matter of course.

10. References

Commonwealth of Australia: *Occupational Health and Safety (Commonwealth Employment) Act 1991*, Australian Government.

Australian DoD (2003): *Navy Technical Regulations Manual ABR 6492*, Australian Department of Defence, July 2003.

Robinson (2003): *Common Law Safety Cases*: Proceedings of the Fifth International Conference on Safety in Road and Rail Tunnels. Marseille, France.

Kelly, T., McDermid, J. (1998): *Safety case patterns-reusing successful arguments*, Understanding Patterns and Their Application to Systems Engineering (Digest No. 1998/308), IEE Colloquium on. London, UK.

Chinneck, P., Pumfrey D., McDermid J.: The HEAT/ACT Preliminary Safety Case: A case study in the use of Goal Structuring Notation. 9th Australian Workshop on Safety Related Programmable Systems. Brisbane, QLD, Australia.