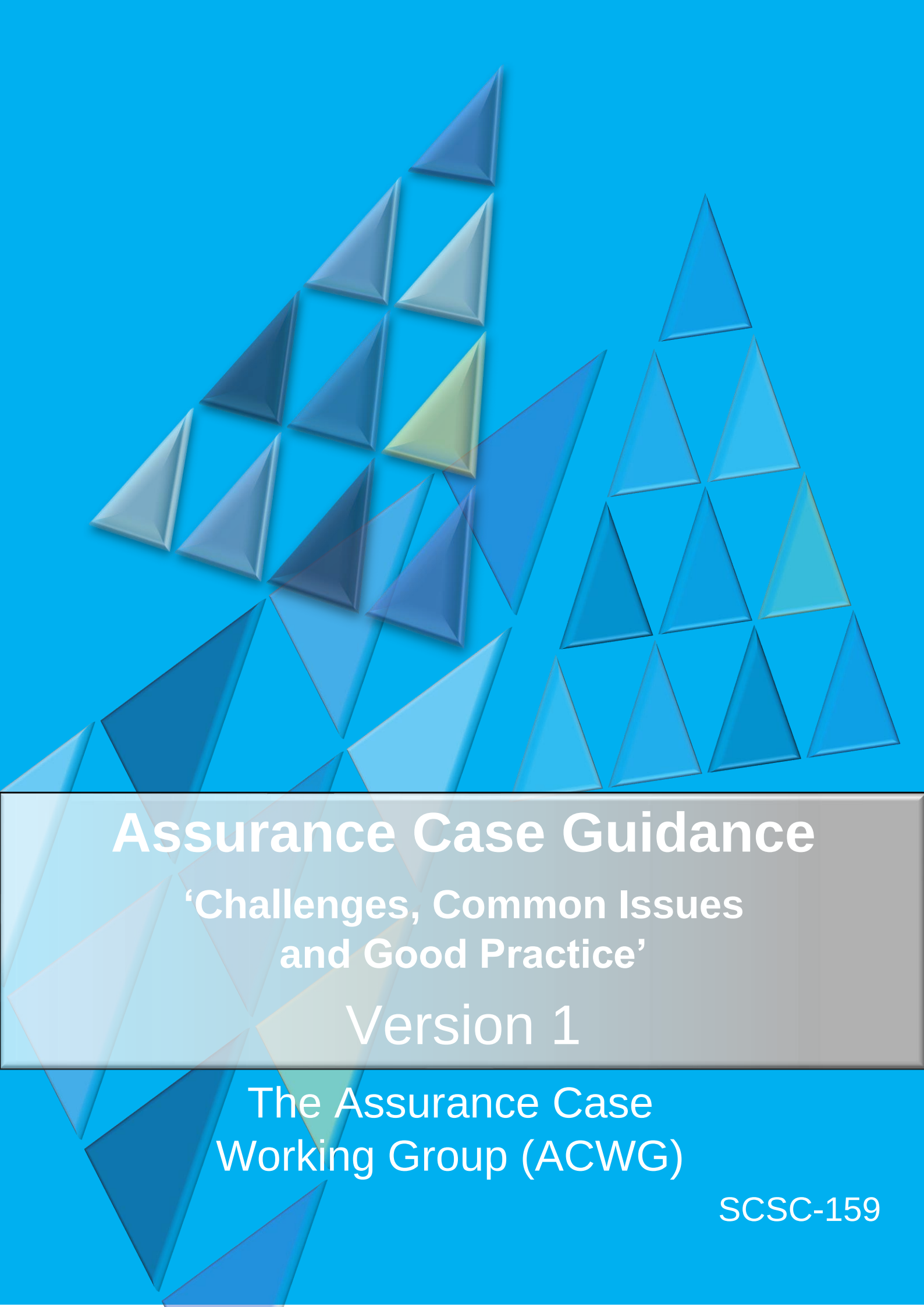


Assurance Case Guidance

‘Challenges, Common Issues and Good Practice’

Version 1

The Assurance Case Working Group (ACWG)

SCSC-159

Assurance Case Guidance

FOREWORD

This guidance has been created to support the production and maintenance of assurance cases that are complete, coherent and proportionate to the confidence required. It is aimed at those who create, review, approve and use assurance cases and is not sector specific.

The guidance focuses on topics that are perceived by the Assurance Case Working Group (ACWG) of the SCSC[1] as containing weaknesses or poor practice, and where no or limited guidance currently exists; it is not intended to be a tutorial. It avoids repeating established guidance; where such guidance exists, it is referenced.

The guidance takes the form of a framework that provides a general introduction to the scope and context of assurance cases supported by papers in Part 2 that each address a specific guidance topic. Part 3 provides supporting information in the form of terminology and references.

The guidance was developed by means of a consensus process involving assurance case authors, reviewers and assessors from academia and industry.

DISCLAIMER

While the authors and the publishers have used reasonable endeavours to ensure that the information and guidance given in this work is correct, all parties must rely on their own skill and judgement when making use of this work and obtain professional or specialist advice before taking, or refraining from, any action on the basis of the content of this work. Neither the authors nor the publishers make any representations or warranties of any kind, express or implied, about the completeness, accuracy, reliability, suitability or availability with respect to such information and guidance for any purpose, and they will not be liable for any loss or damage including without limitation, indirect or consequential loss or damage, or any loss or damage whatsoever (including as a result of negligence) arising out of, or in connection with, the use of this work. The views and opinions expressed in this publication are those of the authors and do not necessarily reflect those of their employers, the SCSC or other organisations.

LICENSE

This document can be accessed for free at scsc.uk/SCSC-159

[1] SCSC : Safety-Critical Systems Club C.I.C. A Community Interest Company registered in England (Company number 13084663)

DOCUMENT HISTORY

Version	Date	Purpose
1	2nd August 2021	Initial Issue

CHANGE HISTORY

This is the first version of the guidance

CONTENTS

PART 1 FRAMEWORK

1:1 BACKGROUND

The SCSC is the UK's professional network for sharing knowledge about safety-critical systems. It brings together engineers and specialists from a range of disciplines working on safety-critical systems in a wide variety of industries, academics researching the arena of safety-critical systems, providers of the tools and services that are needed to develop the systems, and the regulators who oversee safety. It provides opportunities to network and benefit from shared experiences in working hard at the accidents that don't happen; it achieves this through the annual Safety Systems Symposium (SSS) supported by publications, seminars, workshops, tutorials and the collaboration/sharing features of its web site. It focuses on current and emerging practices in safety engineering, software engineering, and product and process safety standards.

Whilst the SCSC is focussed on matters relating to safety, this guidance is intended to be applicable to any property, or combination of properties, of an entity that requires assurance. This guidance was written by the Assurance Case Working Group (ACWG), which is convened under the auspices of the SCSC.

The ACWG is made up of engineers and specialists from a range of backgrounds across industry, consultancy and academia with a wealth of experience in writing, reviewing and using assurance cases. A list of contributors is provided in Section 3:3.

The guidance supports the ACWG's vision:

> *"To produce clear, non sector-specific, guidance on production and maintenance of assurance cases that are complete, coherent and proportionate.*
>
> - *The guidance will focus on topics that are perceived by the ACWG as containing weaknesses or poor practice and where no, or limited, guidance currently exists.*
> - *It will avoid repeating established guidance. Where guidance exists, it will be sign-posted from the ACWG guidance.*
> - *Where there is an opportunity to influence standards and guidance development, the ACWG will provide a focal point for assurance cases."*

In addition, the ACWG maintains the Goal Structuring Notation (GSN) Community Standard (see scsc.uk/gsn), responding to feedback on its use and evolving good practice.

Comments on these documents are actively encouraged. These can be submitted to the Assurance Case Forum scsc.uk/f144 or emailed to the working group using the contact details at scsc.uk/g.

1:2 INTRODUCTION

1:2.1 The need for this guidance

A number of standards have explicit or implicit obligations for assurance of risk-based properties. However, they often contain little guidance on how this should be demonstrated, for example, how to present a case with clarity, or how to determine whether it is sufficiently robust for its purpose. This guidance addresses this shortfall in guidance through an introduction to assurance cases and relevant supporting information.

The underpinning principles of assurance can be applied to any property. The focus of this guidance is on risk-related properties, for example: safety; security; availability.

1:2.2 Structure of the Assurance Case Guidance

This guidance is published as a freely downloadable publication on the SCSC website (see scsc.uk/SCSC-159). This document comprises:

- a framework that introduces the scope and context of assurance cases (Part 1);
- topic papers[2], which each address a specific guidance topic that is perceived by the ACWG as containing weaknesses or poor practice, and where no or limited guidance currently exists (Part 2);
- supporting information including terminology, references and acknowledgements (Part 3).

Guidance topics that support this version are:

1. Avoiding Bias in Assurance Cases

   Assurance cases are occasionally criticised for being biased in the way that they present their argument and evidence. This paper identifies a number of common biases to be aware of (and thus avoid) when constructing, reviewing and using an assurance case.

2. Risk versus Benefit

   This paper provides a discussion on the balancing of risk and benefit. It outlines: the need to balance risk against benefit and how to tell when such an argument may be required. It provides examples where risk and benefit have been successfully balanced, and of the structure for a risk-benefit argument.

3. Modular Assurance

   This paper deals with the structure and presentation of large and complex assurance arguments to support reader comprehension and addresses the benefits in managing the impact of change.

4. The Risk-Confidence-Conformance Approach

   This paper introduces the concept of an approach that facilitates clear presentation of arguments structured around the three central themes of risk, confidence and conformance. It addresses the benefits of the approach and offers examples of how these can be presented.

[2] Part 2 of this document contains papers that each address a specific topic. We refer to them collectively as ‘topic papers’, or individually by their instantiated name (e.g. the Dialectic Argument paper) and cross-reference them by their section number (e.g. ‘the Dialectic Argument paper contained in Section 2:6).

5. Dialectic Arguments in Assurance Cases

   This paper introduces the concept of dialectic arguments and shows how it can be used to add confidence to an assurance argument while it is being authored or during review/evaluation.

Guidance topics in preparation include:

- Proportionality in Assurance Cases
- Assurance Cases across the Supply Chain

Additional papers will be added to future versions of this document as they become available.

1:2.3 Stakeholders

The guidance considers the potential interest of a range of stakeholders without specifying the nature of their interest. Stakeholders considered include product owner, design authority, designer, developer, supplier, operator, user, integrator, auditor, maintainer, regulator, approval authority (internal and external), 3rd party, legal party, insurer, shareholder, customer/procurer, duty holder, etc.

1:3 PURPOSE OF AN ASSURANCE CASE

The primary purpose of an assurance case is to communicate to the relevant stakeholders the assurance argument and supporting evidence for the assurance of the properties of interest of the entity that is the subject of the case.

Notes:

1. The case may address one or more properties of interest and are typically risk based properties, for example: safety; security; availability.
2. The entity that is the subject of the case will depend on the context and scope of responsibility and can include: system; product; component; element; facility; service; activity; process; procedure. Throughout the rest of this document the term *[system]* is used to denote the entity that is the subject of the case. It can be substituted with the relevant term according to context and scope.

The information communicated by the case includes:

- the scope of the case (including the boundary of the *[system]*);
- the acceptability criteria of the performance of the *[system]* with respect to the properties of interest and the confidence in the assurance of such performance, together with their justification;
- the performance with respect to each of the properties of interest of the *[system]* and associated confidence;
- the results of comparing the performance and confidence of the *[system]* with the acceptability criteria;
- any deficiencies, whether in known performance or in the evidence of performance;
- risk(s) associated with the important properties of the *[system]*;
- statements and assumptions about the use and management of the *[system]*, and its operating context;
- a historical record of design decisions and rationales; including any trade-offs that had to be made between the various important attributes;
- information on the limits of the validity of the case.

The assurance case should be used throughout the lifecycle of the *[system]* to inform decisions. For example, the case should inform:

- acceptance of the fundamental concept and safety strategy
- the decision to deploy;
- whether changes are required;
- whether further evidence needs to be collected;
- the decision to retire the *[system]*;
- the prioritisation of actions.

This means that the assurance case needs to be seen as a living and integral part of managing the *[system]*; rather than purely as a deliverable to satisfy a perceived obligation and thereby avoid a *"process-dependent, paper-reliant, 'box-ticking' safety culture"* [1].

The relevant stakeholders may include:

- the developer of the *[system]*;
- the recipients of the *[system]* (e.g. integrators, operators);
- the assessors (e.g. certification bodies, auditors, regulators);
- the affected parties (e.g. members of the public).

It is particularly important that the assurance case is used to inform key decision points in the lifecycle such as:

- safety concept and requirements acceptance
- internal acceptance;
- acceptance by a customer;
- approval by a regulator or similar authority.

1:4 WHAT IS AN ASSURANCE CASE?

An assurance case is defined in [2] as a *"reasoned, auditable artefact created that supports the contention that its top-level claim (or set of claims), is satisfied, including systematic argumentation and its underlying evidence and explicit assumptions that support the claim(s)",* noting that *"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); and
- **justification** of the choice of top-level **claim** and the method of reasoning."

[emphasis added]

An assurance case makes a claim regarding some property and provides arguments, evidence and, where appropriate, assumptions to support[3] it, establishing a conclusion regarding it and the conclusions' associated uncertainty in a given context. It can be used for any property, and is therefore not restricted to 'safety'.

1:4.1 The Structure and Format of Assurance Cases

There is no universally agreed correct structure, format or notation of an assurance case, whether addressing general assurance, or assurance of specific properties, such as safety, of the *[system]*. Sector specific standards may impose requirements for specific content or general arrangement to address the needs of a sector or domain, and to provide consistency for the user of the case.

This guide does not define a particular structure, format or notation for the assurance case but does identify essential elements that should form part of any case, and indicators of good practice.

1:4.2 Good practice in Assurance Cases

Addressing the following topics will increase confidence that the outcome is defensible:

1:4.2.1 Lifecycle

The assurance case should be kept up to date with the *[system]*. It should provide timely updates to inform relevant decisions throughout the lifecycle of the *[system]*, and should be dynamic in response to changes in design, use, context or knowledge arising from use.

Notes:

3. The lifecycle over which the assurance case evolves should include all stages from concept to retirement[4].
4. It may be convenient to separate the case into parts that deal, for example, with: the raison d'être and requirements; implementation; operational use and maintenance.
5. It may be useful to separate the case into elements that divide the *[system]* into elements, potentially with different responsibility or ownership.
6. Where the claim(s) of the case require support from other assurance cases, or the assurance case has been divided into parts/elements, the relationship between the cases/parts/elements should be managed to ensure cohesion and relevance to the scope of the parent assurance case.

[3] It should be noted that bias towards seeking support for the claim can reduce confidence in the conclusions of the assurance case. This concern is addressed further in the Avoiding Bias In Assurance Cases paper (see Section 2:2).

[4] Per [3], the typical system life cycle stages include concept, development, production, utilization, support, and retirement. Other terms may be used for a particular context without changing the intent of this scope.

During assurance case evolution, it is likely the case will establish claims that are not currently supported by evidence. These are aspirational claims and are typically used to provide confidence that there is a strategy for eventual support for such claims. Assurance activities may be established to generate evidence to support the arguments that in turn support these claims. It is essential that a sufficient critical review is included that ensures that these claims do not become self-fulfilling prophecies, and that the assurance activities have sufficient scope to avoid bias and to seek counter-evidence.

1:4.3 Documentation

The baselined assurance case should exist as an artefact[5], or collection of artefacts. These should be controlled in accordance with the organisation's quality management system [4].

Notes:

1. The artefact(s) can exist as a traditional paper or electronic media object(s), or as a model(s) within a tool(s).
2. The artefact(s) need to be maintainable over the life of the subject of the assurance case; particular attention needs to be paid to the through-life supportability of any tools used to create, maintain, publish and distribute the artefacts(s).
3. The artefact(s) should be placed under configuration control so that the relevant version can be clearly identified and differentiated from obsolete/deprecated versions.
4. The artefact(s) should be easily readable and comprehensible by the intended readership:
   - font and graphics used should be sufficiently large and clear;
   - language should be unambiguous;
   - terms, abbreviations and notation should be clearly defined.
5. The artefact(s) should be auditable; claims, evidence, assertions, assumptions etc. should be traceable to their source through explicit, configured references.
6. Some evidence may be difficult to record, for example adding a label to a physical item. In such cases consideration should be given to providing alternative records, for example a record of the process leading to the evidence or a record of its inspection.

As an assurance case can be a significant body of material it is common to provide a summary of the current status within a report often referred to as an assurance case report. This report provides a concise summary of the argument and directs the reader to the full assurance case material.

1:4.3.1 Validity

The assurance case should be valid[6] for its declared scope.

It is noted that claims may not be fully supported when initially recorded in the assurance case (see 1:4.2.1). In this situation an argument may be considered valid but unsupported, and therefore the overall case is not yet sound.

Notes:

1. The scope of the case should be clearly defined and bounded to provide an unambiguous context for the claims and arguments.
2. The case should make one or more claims stated as propositions that are relevant to the subject, property and scope being assured.
3. The arguments should be comprehensive and coherent.
4. The arguments should be valid (see 3:1.2.1).

[5] An artefact can be a document, record, model or other suitable representation of information.
[6] See Assurance Terms (Section 3:1.2.1) for a definition of terms such as 'valid' and 'sound' used in this section.

5. The arguments should be structured with clear and logical decomposition until supported by evidence (see 1:4.3.2).
6. Where an assertion is made without support of evidence, its truth should be self-evident to practitioners within the domain of the *[system]*, within the context of the argument.
7. The arguments should be able to withstand reasonable criticism or objection.

1:4.3.2 Evidence

Evidence that supports the arguments of the assurance case should assert the truth of the premises of the argument.

Notes:

1. Evidence should be relevant to the subject and scope of the case:
   - differences between the evidence basis and the claim context should be justified;
   - assertions drawn from the evidence should be obvious from the evidenced artefacts. Where relevance is not obvious, an evidence assertion can be used to state what is being claimed from the evidence; and support a justification of its relevance.
2. Evidence should be trustworthy:
   - compiled/approved by competent, authorised individuals/organisations;
   - subject to configuration control;
   - verifiable and auditable.

Evidence may not fully support the intent of the original claim or may provide counter-evidence to a claim. Such evidence should not be ignored or excluded; rather the argument and/or claim(s) should be amended to show validity of the claim(s) even in the presence of such evidence. Alternatively the claims should be modified to reflect the uncertainty introduced by the counter-evidence.

1:4.4 Representation of the Argument

Whilst the use of wholly text-based assurance cases may allow distribution to the widest potential audience in a cost-effective manner, they present many weaknesses [8]:

- not everyone is capable of writing in clear, natural language;
- it can take many readings to decipher the meaning of the text;
- multiple cross-references in text can be awkward and disrupt the flow of the main argument;
- it is difficult to develop a clear, shared understanding of the argument.

As such, the use of graphical notations[7] in support of assurance cases are preferred, unless the above weaknesses can be mitigated. It is important that diagrams depicting the argument are appropriate for the target readers of the case, and therefore that an appropriate style of presentation is used.

[7] Graphical notations include Goal Structuring Notation (GSN) [5], Claims-Argument-Evidence (CAE) [6] and Structured Assurance Case Metamodel (SACM) [7].

1:5 GUIDANCE, PRACTICE AND STANDARDS FOR ASSURANCE CASES

Several organisations, associations and international standardisation organisations have issued, or are planning to issue, requirements or guidelines related to an assurance case approach. The following sections contain references to such guidance, practice and standards, organised by domain of application. The referenced artefacts recognise their existence and their inclusion here should not be taken to imply their endorsement.

Where reference is made to a specific property e.g. 'Safety' in the following, the content has been developed to guide the assurance of that property, but can often be applied to other properties.

1:5.1 General

a) Technical Report SRI-CSL-15-01, July 2015 - The Interpretation and Evaluation of Assurance Cases: John Rushby
[http://www.csl.sri.com/users/rushby/papers/sri-csl-15-1-assurance-cases.pdf]
"The first part of this report (Chapters 1–4) provides an introduction to assurance cases. Although this material should be accessible to all those with an interest in these topics, the examples focus on software for airborne systems, traditionally assured using the DO-178C guidelines and its predecessors.
The second part (Chapters 5 and 6) considers the criteria, methods, and tools that may be used to evaluate whether an assurance case provides sufficient confidence that a particular system or service is fit for its intended use"

b) An Overview of John Rushby's Papers on Assurance Cases
[http://www.csl.sri.com/users/rushby/assurance-cases.html]
"[Rushby has] several papers on assurance cases, some of which share text and diagrams, and you might wonder how they relate to each other, or whether [Rushby is] plagiarizing [him]self. [the overview] provides a guide to what [Rushby] thinks are the significant aspects of each paper. The various papers record the evolution of [Rushby's] thinking on the topic of assurance cases."

c) NASA report on effective safety cases
[https://ntrs.nasa.gov/citations/20170003806]
"This report is the result of [a] year-long investigation into assurance case practices and effectiveness … represent[ing] a significant thread of applied assurance methods extending back many decades and being employed in a range of industries and applications. [The] research presented in this report includes a literature survey of over 50 sources and interviews with nearly a dozen practitioners in the field. We have organized our results into seven major claimed assurance case benefits and their supporting mechanisms, evidence, counter-evidence, and caveats."

d) ISO/IEC/IEEE 15026 - Systems and software engineering — Systems and software assurance
Part 1: Concepts and vocabulary (2019)
Part 2: Assurance case (2011)
Part 3: System integrity levels (2015)
Part 4: Assurance in the life cycle (2021)
"The essential concept introduced by ISO/IEC/IEEE 15026 (all parts) is the statement of claims in an assurance case and the support of those claims through argumentation and evidence. These claims are in the context of assurance for properties of systems and software within life cycle processes for the system or software product.
Assurance for a service being operated and managed on an ongoing basis is not covered in ISO/IEC/IEEE 15026 (all parts)."

e) Structured Assurance Case Metamodel (SACM)
Object Management Group (OMG) [https://www.omg.org/spec/SACM]
"This specification defines a metamodel for representing structured assurance cases and a graphical notation for depicting an Assurance Case."

f) Goal Structuring Notation (GSN)
Safety Critical Systems Club – Assurance Case Working Group
[https://scsc.uk/gsn]
"[The Goal Structuring Notation] is a generic argument structuring language, which can be used to document arguments in any domain.
[The GSN] Standard has two intended functions:
Firstly, it seeks to provide a comprehensive, authoritative definition of the Goal Structuring Notation (GSN).
Secondly, it aims to provide clear guidance on current best practice in the use of the notation for those concerned with the development and evaluation of engineering arguments – argument owners, readers, authors and approvers."

g) Claims Argument Evidence (CAE)
Adelard
[https://claimsargumentsevidence.org]
"Claims Arguments Evidence (CAE) is a framework for reasoning and communicating. The initial application was in reasoning to explore the safety and trustworthiness of systems as captured in safety and assurance cases. However, CAE is more broadly applicable to reasoning about complex systems where confidence comes from challenging and developing understanding across disparate groups of people and disciplines. CAE can be used throughout the lifecycle, from brainstorming and optioneering to detailed rigorous explanations."

h) Modular Software Safety Case (MSSC)
Industrial Aviation Working Group (IAWG)
[https://github.com/IAWG/MSSC-Process]
"The Modular Software Safety Case (MSSC) process controls complexity of the argument made for multiple modules of software, potentially across multiple software suppliers, by creating an optimised safety case architecture which reflects both interfaces within the design domain and how likely the design modules are to be subject to future change. This permits composition of the whole safety case through integration of safety case modules of argument whose well-defined interfaces provide some element of resilience to the impact of change"

1:5.2 Medical

a) AAMI TIR38:2019 - Medical device safety assurance case **guidance**
Association for the Advancement of Medical Instrumentation
[https://www.aami.org]
*"This technical information report provides **guidance** on how to complete an Assurance Case Report in order to comply with the new additional [Food and Drug Administration] FDA pre-market requirements for infusion pumps. It includes a detailed but strictly hypothetical example from the medical device domain.*
This TIR provides information useful to creating and maintaining safety assurance cases for medical devices, including drug delivery combination products. It does this in the context of ANSI/AAMI/ISO 14971 and ISO/IEC 15026-2."

b) Infusion Pumps Total Product Life Cycle - Guidance for Industry and FDA Staff
US Food & Drug Administration
[https://www.fda.gov/media/78369/download]
"The Food and Drug Administration (FDA) has developed this guidance document to assist industry in preparing premarket submissions for infusion pumps and to identify device features that manufacturers should address throughout the total product life cycle. Infusion pumps, as described in [Code of Federal Regulations] 21 CFR 880.5725, are intended for use in a health care facility to pump fluids1 into a patient in a controlled manner."

1:5.3 Nuclear

a) NS-TAST-GD-051 Revision 7 - The Purpose, Scope, and Content of Safety Cases
UK Office for Nuclear Regulation (ONR): Nuclear Safety Technical Assessment Guide (TAG)
[https://www.onr.org.uk/operational/tech_asst_guides/ns-tast-gd-051.pdf]
"This technical assessment guide (TAG) for ONR inspectors is on the purpose, scope and content of safety cases. The SAPs use the term 'safety case' "to encompass the totality of the documentation developed by a designer, licensee or duty-holder to demonstrate high standards of nuclear safety and radioactive waste management, and any subset of this documentation that is submitted to the Office for Nuclear Regulation (ONR)". For the purposes of this guidance, the term dutyholder shall be used to refer to any organisation with responsibility for a safety case."

1:5.4 Automotive

a) ANSI/UL4600 - Standard for Evaluation of Autonomous Products
Underwriters Laboratories autonomous vehicle Standard
[https://ul.org/UL4600]
"encompasses fully autonomous systems that move such as self-driving cars along with applications in mining, agriculture, maintenance, and other vehicles including lightweight unmanned aerial vehicles (UAVs).
The Standard uses a claim-based approach which prescribes topics that must be addressed in creating a safety case. It is intended to address changes required from traditional safety practices to accommodate autonomy, such as lack of human operator to take fault mitigation actions.
Topics covered in the Standard include safety case construction, risk analysis, safety relevant aspects of design process, testing, tool qualification, autonomy validation, data integrity, human-machine interaction (for non-drivers), life cycle concerns, metrics and conformance assessment. Security is addressed as a requirement, but details are currently outside the scope of the proposed Standard."

b) ISO 26262 - Road vehicles - Functional safety
Part 2: Management of functional safety (2018)
Part 8: Supporting processes (2018)
Part 10: Guidelines on ISO 26262 (2018)
"A safety case shall be developed, in accordance with the safety plan, in order to provide the argument for the achievement of functional safety.
The safety case should progressively compile the work products that are generated during the safety lifecycle to support the safety argument.
In the case of a distributed development, the safety case of the item can be a combination of the safety cases of the customer and of the suppliers, which references evidence from the work products generated by the respective parties. Then the overall argument of the item is supported by arguments from all parties. The interfaces between the customer and a supplier are defined in a Development Interface Agreement
To support safety planning, the intended safety arguments can be identified prior to work products becoming available. To support progressive functional safety assessments the safety case can be released progressively as work products are generated to provide evidence for the safety arguments."

1:5.5 Rail

a) EN 50129 (2018) Railway applications - Communication, signalling and processing systems - Safety related electronic systems for signalling
"This standard defines requirements for acceptance of safety-related electronic systems in the railway signalling domain, and it prescribes the structure and content of the safety case, which contains evidence of quality management, of safety management, and of functional and technical safety. The safety case consists of 6 parts, i.e. (1) definition of system, (2) quality management report, (3) safety management report, (4) technical safety report, (5) related safety case and (6) conclusion. Moreover, (4) technical safety report is formed by 6 sections, i.e. (a) introduction, (b) assurance of correct functional operation, (c) effects of faults, (d) operation with external influences, (e) safety-related application conditions and (f) safety qualification test results. In addition, the development of a railway signalling system can be classified into three types, i.e. (i) generic product, (ii) generic application and (iii) specific application, and three different categories of safety case are defined according to these development types."

1:5.6 Air

a) Guidance Material on Building a Safety Case for Delivery of an ADS-B Separation Service
International Civil Aviation Organization (ICAO)
https://www.icao.int/APAC/Documents/edocs/cns/APX. J - Guidance Material on Building Safety Case for ADS-B separation.pdf
"Basic guidance on the building of a Safety Case for delivery of an ADS-B separation service is provided in this document.
A number of discrete 'steps' in the building of a Safety Case are described to progress to a completed document. The first steps cover the generic requirements for the preparation of a Safety Case for any airways system, including any surveillance systems used for separation by ATC. The remaining steps cover the elements of a Safety Case specific to a new ADS-B surveillance service. The final steps are guidance to the actual content headings of a Safety Case for an ADS-B service."

1:5.7 Oil and Gas

a) Safety cases
UK Health & Safety Executive (HSE) > Guidance> Industries> Offshore oil and gas> Offshore topics
[https://www.hse.gov.uk/offshore/safetycases.htm]
"This site helps dutyholders to comply with the legal requirements, and shows HSE's procedures for handling safety cases and other submissions."

b) Safety Cases - Competent Authority Portal ('CAP') : Industry user guidance
UK HSE> Guidance> Industries> Offshore oil and gas> Offshore Safety Directive Regulator (OSDR)> OSDR Guidance
[https://www.hse.gov.uk/osdr/safety-cases/index.htm]
"The Offshore Safety Directive Regulator (OSDR) [has] develop[ed] an online portal for the submission of documents by the Industry and for the [Competent Authority] to carry out its functions as required under the Directive [on the safety of offshore oil and gas operations]."

c) Submitting gas transporter safety cases - Gas Safety (Management) Regulations 1996 (GSMR)
UK HSE> Guidance> Industries> Gas> Gas supply> Legislation
[https://www.hse.gov.uk/gas/supply/gasscham/index.htm]
"…a safety case has to be formally accepted by HSE before [a] dutyholder can convey gas. Safety cases should be prepared according to the requirements of GSMR and … the guidance on GSMR (L80) and with reference to the Safety Case Assessment Manual.
The purpose of the manual is to provide HSE assessors with guidance on assessing and accepting a safety case. It also provides transparency to dutyholders about the standards and timescales that HSE uses and promotes consistency in the judgements made in the assessment process."

d) Safety cases and validation
Australia National Offshore Petroleum Safety and Environmental Management Authority (NOPSEMA) Home> Offshore industry> Safety> Safety cases and validation
[*https://www.nopsema.gov.au/offshore-industry/safety/safety-cases-and-validation*]
"*A facility cannot be constructed, installed, operated, modified or decommissioned without a safety case in force for that stage in the life of the facility.*
The operator of a facility must submit the safety case to NOPSEMA with either a NOPSEMA pro-forma cover sheet or a covering letter stating that it is being submitted for assessment. Since it is the operator that must submit the safety case, registration of the operator must be completed (and a scope of validation for a proposed facility agreed) prior to safety case submission.
The Offshore Petroleum and Greenhouse Gas Storage (Safety) Regulations 2009 (OPGGS(S)) set out the requirements for the contents of safety cases. These Regulations state that NOPSEMA may, by notice in writing, require the operator of a proposed facility or an existing facility, to provide a validation in respect of the proposed facility or in respect of a proposed significant change to an existing facility."

1:5.8 Defence

a) SMP12. Safety Case and Safety Case Report
UK Ministry of Defence (MOD) Acquisition Safety & Environmental Management System (ASEMS)
[https://www.asems.mod.uk/guidance/posms/smp12]
"This procedure provides guidance on the development of a Safety Case and Safety Case Report. The Safety Case brings together all project safety information, forming a number of arguments which are summarised in the Safety Case Report."

b) MOD Exemplar Gas Safety Case
UK Ministry of Defence (MOD)
[https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/262274/20131025_MOD_ExemplarGasSafetyCase.pdf]
"MOD, as a Gas Transporter within Great Britain, has submitted this Exemplar Safety Case to demonstrate compliance with the Gas Safety (Management) Regulations 1996."

c) Implementation of improved Air System Safety Case Regulation - RA 1205 [News Item]
UK MOD / Military Airworthiness Authority (MAA)
[https://www.gov.uk/government/news/implementation-of-improved-air-system-safety-case-regulation-ra-1205]
"The MAA recognised an important opportunity to enhance both the activity associated with the in-service ownership and management of [Air System Safety Case] ASSCs, and the effective influence of Air Safety requirements on capability design/selection. Subsequently Niteworks was contracted to investigate the links between the MOD's capability development process and the establishment of effective ASSCs. The resulting study led to a revision of Regulatory Article (RA) 1205 and the introduction of a new Manual of Air System Safety Cases - both documents to be published in a Notice of Proposed Amendment in January 2019."

d) Manual of air system safety cases (MASSC)
UK MOD / Military Airworthiness Authority
[https://www.gov.uk/government/publications/manual-of-air-system-safety-cases-massc]
"The MASSC provides guidance to those organisations required to establish and maintain an Air System Safety Case (ASSC)"

e) DSA 03-OME Part 1 (JSP 520)- Defence Code of Practice (DCOP) and Guidance Notes for [Ordinance, Munitions, Explosives] OME Acquisition Chapter 9: safety and environment case development
UK MOD Defence Safety Authority
[https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/892546/DSA_03_JSP_520_CHAPTER_09__Apr_20_A1_.pdf]

f) DSA03.DLSR.LSSR Land System Safety and Environmental Protection Defence Codes of Practice (DCoP) (Previously JSP 454 Part 2)
UK MOD Defence Safety Authority
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/646340/DSA03-Guidance-DLSR-LSSR.pdf
"The Ministry of Defence (MOD) has a duty to protect its employees, those that may be affected by its activities and the environment. Effective safety and environmental protection is crucial to force protection and maximising operational capability. These DCoPs provide practical advice on how to comply with a Regulation articulated in DSA02.DLSR.LSSR… Guidance material is also included in this document which, whilst not compulsory, may also be considered 'good practice' to further support the Regulations and DCoPs." [Includes guidance on Safety and Environmental Case Development (DCoP D), including advice on proportionality in the development of the case]

PART 2 TOPIC PAPERS

2:1 INTRODUCTION TO THE TOPIC PAPERS

2:1.1 How To Use The Topic Papers

This part contains topic papers[8], which each address a specific guidance topic that is perceived by the ACWG as containing weaknesses or poor practice, and where no or limited guidance currently exists. They are not intended to provide an end-to-end process for assurance case construction, review or maintenance, rather they are intended to be taken to provide insight into typical issues encountered with assurance cases and offer good practice to avoid those issues.

2:1.2 Intersection Of Topics Covered By The Papers

While each paper focusses on a particular topic, the topics overlap. The guidance contained in each may be taken together to inform an adopted approach to assurance case development, review and maintenance, alongside other in-house or industry practice. For example, the use of a dialectic approach as discussed in Section 2:6 may be used to address bias identified in Section 2:2. Similarly, the use of modularised approach to assurance cases as discussed in Section 2:4 may be used to facilitate the structuring of an argument built by consideration of Risk, Confidence and Conformance as addressed in Section 2:5.

2:2 AVOIDING BIAS IN ASSURANCE CASES

2:2.1 Introduction

2:2.1.1 Problem Statement

Assurance cases are occasionally criticised for being biased in the way that they present their argument and evidence [1][9], and this criticism may very well be justified. Assurance cases do not, typically, present balanced arguments in the same way as one would expect to see in a hearing in a court of Law. That is, assurance cases rarely, overtly, contain an argument 'for' and 'against' the claim being made (e.g. they would typically instead argue only that 'System A is safe', or 'System B is secure').

This all makes perfect sense, after all, it's a case that positively supports the demonstration of some *[system]*'s characteristic that stakeholders are looking for (isn't it?). Well yes, ultimately, but perhaps what stakeholders actually need at any time is a more comprehensive, well-balanced (un-biased) and structured account of pertinent information that will enable them to determine an *[system]*'s ability to perform safely, or securely, or reliably etc.

2:2.1.2 Scope

The guidance provided within this paper is intended to be of interest to stakeholders across all sectors, and have applicability to all types of assurance case. For this reason the specific types of bias that are mentioned within the paper are those thought to have broad applicability rather than those that might be seen more frequently within one sector, or impact only one type of analysis or measurement technique or design etc.

[8] Part 2 of this document contains papers that each address a specific topic. We refer to them collectively as 'topic papers', or individually by their instantiated name (e.g. the Dialectic Argument paper) and cross-reference them by their section number (e.g. 'the Dialectic Argument paper contained in Section 2:6).

This paper provides links to other, more detailed, sources of information for those readers who require a greater understanding of the subject and how it might impact their own sector or working practices.

2:2.1.3 Structure

The desire to produce positively-biased assurance cases is now deep-rooted, and any change to this behaviour is likely to require two fundamental changes to the way we think about assurance cases. The first of these changes is relatively straight forward; the purpose (or definition) of an 'assurance case' needs to express a requirement clearly for it to describe both the positive and negative aspects of the *[system]* and property to be assured. Then, secondly, assurance case stakeholders (e.g. authors, reviewers, duty holders and regulators) need to be aware of, and therefore introduce controls to manage, their own cognitive biases when creating, reviewing or drawing conclusions from the case being presented.

This paper discusses each of these issues in turn. It provides examples of how they might, if not addressed, introduce bias into the assurance case creation or review effort. Then, by proposing a number of changes to current, common practice, it highlights opportunities for improvement that might go some way towards addressing the criticisms of bias in assurance cases.

2:2.2 Definition and Purpose: An Inextricable Link

2:2.2.1 The Problem with Existing Definitions

When reviewing existing assurance cases for 'fitness for purpose' it is often easy to identify weaknesses, especially when they are scrutinised in the wake of significant events, e.g. large scale accidents. The presence of bias, hitherto undetected, is one characteristic of an assurance case that is often exposed through post-accident/incident inquiry. However, before we ourselves point fingers and sneer it is worth reminding ourselves, if only briefly, of the purpose of an assurance case. Scott and Krombolz [10] define a structured assurance case as follows:

> *A documented body of evidence that provides a convincing and valid argument that a specified set of critical claims regarding a system's properties are adequately justified for a given application in a given environment.*

The UK Military Aviation Authority (MAA) [11] defines an assurance case (for safety) as:

> *A structured argument, supported by a body of evidence that provides a compelling, comprehensible and valid case that an air system is safe for a given application in a given environment. It is throughlife, pan-Defence Lines of Development (DLoD) and addresses a combination of the physical components, procedures and human resources organized to deliver the capability.*

The UK Civil Aviation Authority (CAA) [12] defines an assurance case (for safety) as:

> *A document which clearly and comprehensively presents sufficient arguments, evidence and assumptions that system hazards have been identified and controlled for both engineering and operational areas to demonstrate that a facility, facilities or organisation is/are adequately safe in air traffic service requests.*

These are all fairly standard definitions; what they have in common is that they reinforce the perception that the purpose of an assurance case is primarily to demonstrate the achievement of some desirable attribute, for example, that System A is acceptably safe when used within environment Y. Haddon-Cave, the author of the Nimrod Review [1], criticised this perception when he stated that:

> *Safety cases tend also to give the answer that the customer or designer wants, i.e. 'that the platform is safe'.*

This criticism can be levelled at any kind of assurance case. The implication within Haddon-Cave's comment is that assurance cases should, in fact, support an objective assessment of some characteristic of an *[system]*; that is, they should, where appropriate to do so, contain both confirmatory and non-confirming arguments and evidence, thereby allowing users to draw conclusions based upon a more complete and balanced understanding of the various influencing factors.

The UK Office for Nuclear Regulation (ONR) [13] defines an assurance case as follows:

> *A safety case is a logical and hierarchical set of documents that describes risk in terms of the hazards presented by the facility, site and the modes of operation, including potential faults and accidents, and those reasonably practicable measures that need to be implemented to prevent or minimise harm. It takes account of experience from the past, is written in the present, and sets expectations and guidance for the processes that should operate in the future if the hazards are to be controlled successfully. The safety case clearly sets out the trail from safety claims through arguments to evidence*

This definition differs slightly from the others, and it appears to go some way towards addressing Haddon-Cave's criticisms. For example, according to the ONR's definition an assurance case is required to describe *the risk presented by a facility, site and the modes of operation, and those reasonably practicable measures that need to be implemented to prevent or minimise harm*; it is not, apparently, required to state that the facility, site or mode of operation is safe (this is presumably the responsibility of the associated license holder having reviewed the content of the assurance case). This is an important distinction that may help identify good practice.

Good practice may also, perhaps, be identified by considering the use of 'cases' in Law. There, the case for the prosecution is different to, and separate from, the case for the defence; however, both cases are presented and cross examined before the judge and jury for consideration before decisions are made. If the legal model is adopted then perhaps it is perfectly acceptable for an assurance case only to argue that an *[system]* is 'acceptably safe/secure etc.'; perhaps what is missing is a separate case that (attempts to) demonstrate the opposite, that the *[system]* is unsafe, or insecure etc.

The problem often encountered today is that, notwithstanding the significant effort that goes into generating evidence to underpin an assurance case, often it is only the positive (supporting or /confirmatory) evidence that is collated, analysed and structured in such a way as to support a compelling argument that some similarly positive claim is valid. The evidence that detracts from, or undermines, the demonstration of a positive high-level claim may not be pursued or analysed to the same degree, and is rarely utilised to construct an explicit and robust argument for the 'negative' claim.

Whether there are two distinct cases (e.g. a 'safety case' and an 'unsafe case') or just one case that applies proportionate effort to arguing both viewpoints is unimportant; a satisfactory conclusion could be drawn no matter which approach is adopted. What is important is the need to understand that, if the final conclusions drawn from analysing the content of an assurance case are to be robust, i.e. capable of withstanding scrutiny (perhaps in a court of Law), the 'negative' or 'non-confirming' case' must exist, in some form, and the case it presents must be considered by the Duty Holder alongside that of the 'positive' or confirming' assurance case'.

2:2.2.2 Must an Assurance Case Demonstrate that a *[system]* is assured?

The preceding section raises an interesting question that is worthy of further discussion: must an assurance case demonstrate satisfaction of some high-level claim in order to be considered successful?

Demonstrating that a *[system]* is sufficiently safe, or secure, or reliable will almost certainly be an aspiration for stakeholders; but, must an assurance case demonstrate that its high-level claim has been achieved before it can be judged valid or successful? When one examines the definitions given above it is made quite clear that this 'satisfied' status *is* often a requirement; so, if an assurance case cannot, at some point in time, demonstrate achievement of the high-level claim is it no longer a valid assurance case? And if not, what is it, and does it still have any value?

These questions may appear trivial, but the point being made is not. An assurance case that cannot demonstrate achievement of its top-level claim *is* still a valid assurance case. More than this, if one assumes that it is well-structured and evidenced then it is also a successful assurance case. This is because it is effective in relaying the message that the *[system]* does not yet achieve a particular high-level objective; for example, it may undermine a claim that System A is safe – and that is a very important message to get across.

An assurance case does not, therefore, need to demonstrate achievement of some high-level claim in order to be successful or of value. It is equally useful (perhaps even more useful) when it informs stakeholders that some high-level claim (e.g. System A is safe) cannot yet be demonstrated.

2:2.2.3 Redefining the Purpose of an Assurance Case

From the discussion presented above it is apparent that many of the definitions in use throughout the safety-critical systems sector reinforce the perception that the purpose of an assurance case is purely to demonstrate the achievement of some high-level and desirable characteristic, e.g. that a system is 'acceptably safe'. The text also suggests that definitions of this type may be contributing to the introduction of bias within assurance cases.

A simple solution to this problem may be to revise current definitions to ensure they explicitly require consideration of those elements that challenge the thrust of the positive or affirming argument. Leveson [9] stated:

> *To avoid confirmation bias and compliance-only exercises, certification should focus not on showing that the system is safe but in attempting to show that it is unsafe.*

Leveson's proposal seems to support the idea that assurance cases should consider the negative or non-affirming evidence associated with a *[system]*; however, redefining the purpose of an assurance case to ensure that it focusses upon only the negative aspects of a *[system]* might result in an approach that still suffers from bias – only now the problem could be negative bias rather than positive bias.

Perhaps a more appropriate definition for an assurance case would be:

> *A structured argument, supported by a body of evidence that, without bias, makes a case for both the 'safe' and 'unsafe' characteristics of a system when used for a given application in a given operating environment.*

Clearly this example is directed towards *safety* assurance; however, the general wording is equally applicable to other desirable characteristics, e.g. security and dependability.

Intuitively, a definition such as this is more likely to result in the creation of less-biased assurance cases, however, it is recognised that making a change such as this will have an

associated cost. Modifying a definition is simple enough, but significant resource may be required if an organisation is to revisit the assurance cases for all entities within its area of responsibility.

Redefining the purpose of assurance cases is only half the battle. Once we have in place a firm requirement for authors to create well-balanced arguments that seek out positive (affirming) and negative (non-affirming) evidence with equal vigour, it will be necessary to overcome the more personal and often completely unintentional biases that could continue to restrict the development of well-balanced assurance cases – the cognitive biases.

2:2.3 Avoiding (Cognitive) Bias

In response to the issues identified within the Nimrod Review [1], including that cited in Section 2:2.2.1, Haddon-Cave made a number of recommendations to improve working practices and help prevent reoccurrence. The following recommendation is of interest to this paper:

> *Care should be taken when utilizing techniques such as 'Goal Structuring Notation' or 'Claims-Arguments-Evidence' to avoid falling into the trap of assuming the conclusion ('the platform is safe'), or looking for supporting evidence for the conclusion instead of carrying out a proper analysis of risk.*

Clearly Haddon-Cave had similar concerns to those expressed by Leveson, and to those addressed by this paper. The tendency for people to assume outcomes based upon their own experiences, focus on one possibility only and to ignore contradictory or non-confirmatory evidence, i.e. to be influenced by cognitive biases, is a real issue, but it is often the consequence of unconscious behaviour, and it can, therefore, be prevented (at least partially) through education and process.

'Cognitive bias' refers to a systematic deviation from rational thinking when making judgements, such that conclusions (e.g. inferences about people and/or situations) are drawn in an illogical manner. When cognitive bias is at play individuals create a subjective reality based upon their own perception of some 'thing' or 'someone', often unconsciously overlooking the facts embedded within available objective evidence. As a consequence of this, cognitive bias can lead to inaccurate judgment, illogical interpretation or what is commonly termed 'irrational behaviour'.

Research has identified a great number of cognitive biases, and many are capable of influencing the judgment of the various stakeholders to an assurance case. However, the purpose of this text is simply to raise awareness of cognitive biases, and it does this only by introducing the concept at its highest level. It is recommended that those charged with commissioning, authoring, reviewing or relying upon the conclusions drawn within assurance cases refer to the more authoritative texts that have been produced including [14] and [15].

2:2.3.1 Confirmation Bias

Confirmation bias is the tendency to search for, interpret, focus on and remember information in a way that confirms rather than falsifies one's preconceptions. For example, confirmation bias will cause assurance case authors to favour information that confirms previously existing beliefs (or biases) and to discount or play-down information that challenges those perceptions. Unfortunately, confirmation bias can prevent us from assessing situations or information objectively. It can influence the conclusions we reach and lead to poor choices. Assurance case authors must learn to objectively consider the value of all relevant information, even when (perhaps most importantly when) it disconfirms the general thrust of the argument being composed.

2:2.3.2 Disconfirmation Bias

Related to confirmation bias is ‘disconfirmation bias’, the tendency for people to apply critical scrutiny to information or evidence that undermines their prior beliefs, and to accept without scrutiny that information or evidence which is aligned with their prior beliefs. Stakeholders to an assurance case must consciously try to overcome this bias by being open to all information, evidence and viewpoints, even when it contradicts that which they believe to be true. All evidence should equally be subject to challenges of integrity, sufficiency and/or accuracy.

2:2.3.3 Observer-Expectancy Effect

The observer-expectancy effect is when an individual expects a given result and therefore unconsciously misinterprets or misrepresents data in order to achieve it. Assurance case authors in particular must, at all times, be alert to this bias when assessing the true value of the information at their disposal. Though potentially an onerous task, assurance case authors must review all evidence in detail, and reliance must not be placed upon the originator’s conclusions alone.

2:2.3.4 Ostrich Effect

The ostrich effect finds our assurance case stakeholders with their heads in the sand, ignoring an obvious (negative) situation or item of non-confirming evidence. This bias can, in the extreme, cause individuals to unconsciously ignore the fact that non-confirming evidence exists, and they therefore fail to even consider its importance.

2:2.3.5 Conservatism Bias

Conservatism bias is the tendency to reject new evidence or new information because it contradicts one’s established beliefs (i.e. to favour prior evidence over emerging evidence). When considering assurance case authors who have spent months, if not years, creating a robust assurance case for a *[system]*, conservatism bias might see them reject new evidence that calls into questions the validity or sufficiency of some part of the underlying evidence. For example, the suggestion that an increasing number of in-service failures may call into question the conclusions drawn by a related quantitative analysis may be rejected without proper consideration, even though the suggestion is valid and worthy of investigation.

2:2.3.6 Summary

The biases listed above are merely intended to provide examples of the many that exist, and it is acknowledged that they are not necessarily completely distinct, easily identifiable or free from overlapping concepts. They have been included to show how they can easily, and unconsciously, affect decision making amongst assurance case stakeholders. The list is far from exhaustive, it is intended only to raise awareness of the issue, and in doing so, prompt further reading to help ensure that stakeholders introduce appropriate controls to reduce the likelihood of cognitive biases influencing their behaviour during the preparation and/or review of assurance cases.

An important condition associated with all cognitive biases is that they are unintentional or automatic. Clearly, if one deliberately chooses to overlook contradictory or non-supporting evidence then this is the effect of a completely different human trait.

2:2.4 Cultural Changes To Help Avoid Bias in Assurance Cases

The information referenced throughout this topic paper does not, typically, provide solutions or reliable approaches to bias mitigation. Indeed, according to Kahneman [16] the best we can hope to do is learn to recognize situations in which mistakes are likely to be made, and try harder to avoid specific errors when the stakes are high. However, to help avoid the influence

of bias when writing, reviewing or reasoning about assurance cases, organisations may wish to start by encouraging the following behaviours:

Encourage a Hazard/Threat Seeking Culture. Throughout all phases of a *[system]*'s lifecycle, stakeholders should be encouraged to actively seek out and highlight hazards/threats etc. An organisation's senior leadership team should actively promote this type of behaviour and openly reject actions that seek to suppress the identification, and therefore management, of hazards/threats etc.

Seek Disagreement. Organisations should foster an environment where it is not only 'okay' to disagree but also encouraged. Asking colleagues 'Am I right?' will typically attract responses only from those who agree with you. A better question is perhaps 'Why am I wrong?' Organisations should encourage those present during meetings to play devil's advocate in order to test theories and solutions.

Seek, include and discuss counter-evidence. When preparing an assurance case organisations should seek out, include and discuss disconfirming (counter) evidence. Assurance case practitioners should be encouraged (by process) to include counter-evidence within their cases thereby creating more robust arguments and controlling the influence of confirmation bias.

2:2.5 Representing Dialectic Arguments

When selecting a method, technique or notation to describe an assurance case argument, organisations should give some consideration to that approach's ability to represent dialectic arguments (see Section 2:6). If the organisation's intent, as supported by this paper, is to explicitly include counter-argument and counter-evidence within its assurance cases, then the use of a method and/or notation that has been designed to accommodate dialectic arguments is likely to result in cases that are easier to produce, maintain and understand.

2:2.6 Conclusions

This paper discusses the perceived problem of bias within assurance cases. Two potential causes of bias are discussed; the first is a fundamental flaw in the way the principal engineering sectors define the purpose of an assurance case, and the second is the more insidious influence of cognitive bias. The paper proposes an alternative definition for an assurance case that explicitly requires the presentation of a more balanced argument, i.e. one that expends proportionate effort in trying to identify and record the goal-affirming (safe, secure etc.) characteristics of a *[system]* as it does the goal-challenging (unsafe, insecure etc.) characteristics of a *[system]*. The paper also attempts to raise awareness of how cognitive biases can unwittingly, and detrimentally, affect the judgment of the various assurance case stakeholders. It is hoped that by raising awareness of this unintentional, but potentially misleading, human trait those commissioning, creating or interpreting assurance cases will ensure that appropriate and sufficient controls are in place to reduce their likelihood of occurrence.

2:3 RISK VERSUS BENEFIT

2:3.1 Introduction

2:3.1.1 Problem Statement

Every activity has a probability of occurrence of the primary goal(s) and also the secondary effect(s) of the activity. Where these outcomes are positive or as intended, then these can be considered as a success or a benefit. The mirror of this is where the outcomes are negative; in this case the activity may be considered a failure or having caused harm. Most activities are complex and as a result there is a mixture of multiple positive and negative outcomes.

The decision whether to conduct a given activity is often made on the delivery of the primary goal. However, in activities relating to safety, other outcomes are often considered before proceeding. This decision is usually framed in the context of:-

"if I do this activity what are the potential consequences of an in service or product failure?"

What would be the outcome if we considered the wider implications i.e. as well considering the consequences of service failure we considered the positive outcomes in the decision to proceed?

The widely known concept of "As Low As Reasonably Practicable" (ALARP) is a specialised case of risk versus benefit where the potential for harm 'risk' is considered against the time, effort and costs in reducing or removing the risk.

It is assumed for this section that an intolerable risk would result in an activity being stopped, or in the case of a change programme, the implementation of the change suspended. The activity could be restarted or the change delivered once the risk has been reduced to a tolerable level. A risk-benefit argument based upon the framework in Section 2:3.6 provides a method to record the rationale for the acceptance of a tolerable level of risk.

2:3.1.2 Scope

This paper provides a discussion on the balancing of risk and benefit. It is left to the user to determine its applicability to the assurance case they are making. Further information can be found, for example, on the UK Health and Safety Executive website [17].

2:3.1.3 Structure

This paper outlines the need to balance risk and benefit (Section 2:3.3); how to tell when such an argument may be required (Section 2:3.4); examples where risk and benefit have been successfully balanced in Section 2:3.5; and the structure for a risk-benefit argument (Section 2:3.6).

2:3.2 Good Practice

A risk-benefit argument needs to be clear and transparent. This guidance paper provides a framework that could be used for the duty holder to fully understand the options being presented and trade-off that between risk and benefit.

2:3.3 Why Do We Need to Balance Risk Versus Benefit?

Risk is often simply defined as the product of the severity of an outcome and the likelihood of its occurrence. There is no single measure of acceptability for risk, with the units and "quantity" of risk that can be accepted varying from domain to domain.

Where there are no agreed criteria for what an acceptable risk is, then the responsibility for determining the thresholds of acceptability fall upon the risk owner for the activity. This must

be undertaken in a manner which respects the rights of the subjects exposed to the risk to have a reasonable expectation of freedom from harm. This is encapsulated in the Nuremberg Code [18] for clinical research.

Broadly speaking the risk can be divided into three categories:-

1. Minimal Risk[9], where the magnitude of harm is broadly acceptable;
2. Minor increase over minimal risk; and
3. Excessive risk.

These are taken from the medical clinical trial model and are conceptually equivalent to the acceptable, tolerable, and unacceptable regions of the ALARP triangle. The principle is that the risk should always be driven to the minimum practicable level.

It is not always possible or practical for every life activity to be conducted in the acceptable or minimal risk zone. Conversely it is generally good practice to avoid conducting activities in the excessive or unacceptable risk zones. It is operating in the middle region where there is a minor increase over minimal risk that the risk versus benefit argument is required to demonstrate that the activity is worthwhile. This is consistent with the sixth Nuremberg Code principle that

> *"the degree of risk to be taken should never exceed that determined by humanitarian importance of the problem to be solved by the experiment".*

The decision to accept a given risk should be taken by the duty holder[10] for the service or function being provided.

2:3.4 How Do We Determine That a Risk-Benefit Argument is Required?

To understand the risk from an activity, and its acceptability, conventionally a set of assurance requirements are derived by a set of process steps which simplistically could be considered as below:

1) Find the hazards.
2) Identify the severity of the outcomes associated with the hazards.
3) Determine the tolerable likelihood for the given severity of outcome, for the risk to be accepted.
4) Capture the findings of the above as assurance requirements.

The intention of good design is that all assurance requirements will be satisfied, thereby ensuring that the level of risk achieved is at the minimal level. Where this is the case, there is an 'automatic' assumption that the risk is acceptable.

Safety requirements, for example, can be set with a desire to improve the safety of the activity. The desire to improve can be driven by business need, regulation, legislation, or management direction.

If the assurance requirements cannot be met, then it should be asked whether changes can be made so that the assurance requirements can be satisfied. Examples of how this could be achieved include a redesign of key components or a change of use to reduce the severity of the outcome.

---

[9] A risk is minimal where the probability and magnitude of harm or discomfort anticipated in the proposed research are not greater, in and of themselves, than those ordinarily encountered in daily lives of the general population or during the performance of routine physical or psychological examinations or tests.

[10] The 'duty holder' role is a concept frequently adopted to refer to the person or organisation that is accountable in relation to a statutory duty. See glossary Section 3:1.2.4 for further details.

If, after additional work, the assurance requirements cannot be met and the duty holder wishes to consider proceeding, then there needs to be a risk-benefit argument produced justifying that any risk above the minimal level is not excessive and is either:

a) temporary, and will be offset by a future improvement; or

b) permanent, and has other beneficial consequences.

When safety requirements are not satisfied and the residual risk is determined to be above the minimal level, and when subsequent analysis indicates that the benefits do not justify the risk then the activity lies in the 'excessive' or 'unacceptable' risk zones.

When this is the case it is preferable to seek an alternative solution; however, a decision to accept an 'excessive' or 'unacceptable' risk (for the benefit it delivers) can be taken by the duty holder for the service or function being provided.

2:3.5 How Do We Argue Balance of Risks?

Throughout the world there are activities where positive outcome is offset against negative outcome. The examples in this paper illustrate a range of possibilities from where the risk-benefit argument considers the trade-off:

- in potential harm to the same person (see 2:3.5.1);
- through risk transference where one party benefits by another party being exposed to increased risk (see 2:3.5.2); and lastly,
- to the benefit being delivered assessed as to whether the cost is grossly disproportionate (see 2:3.5.3).

Comparisons of risk and benefit are difficult, generally because the risk and the benefit are not measured in common units even when measuring what appear to be related items. Risk of harm to an individual can be measured as estimated fatalities or injuries in a year, or the probability of an aircraft crash with an assumed number of passengers, etc. Benefits could be measured in estimated lives saved, tonnes of carbon-dioxide not produced, cost savings etc. These require a very subjective view and a good argument to illustrate the benefits outweighs the risks.

When comparisons are made, it is always clearest when there are direct quantitative units of the same value being compared. This makes the trading of risk relatively trivial. The example given below for airbags is one such area where this type of argument has been successfully made.

2:3.5.1 Airbags in Cars

Historically, car crashes included cases where there were driver injuries and fatalities caused by the impact of the driver on the steering wheel. The solution to this was a range of measures including fitting an airbag to the wheel. The airbag would deploy in the event of a frontal impact providing a cushioned surface which mitigates the effects of the impact by reduction of the force causing the deceleration.

Airbags were first commonly fitted in the USA [19]. Airbags in the USA are larger and inflate faster than those in the UK because the wearing of seat belts is less common in the USA and they have to provide protection faster. The Royal Society for the Prevention of Accident (RoSPA) reported [19] that the effectiveness of airbags was assessed by the National Highway Traffic Safety Administration (NHTSA) in the USA at an early stage and they estimated that while 4,000 car occupants had been saved by airbags, there were 60 fatalities that were attributed to the airbag, which would have been an otherwise survivable accident. After 30 years of airbag use from 1987 to 2017, the NHTSA declared frontal airbags had saved 50,457 lives, and that improvements in airbag technology had reduced the fatalities from airbags for otherwise survivable accidents [20]. There is also evidence that airbags cause other injuries such as broken limbs, hearing problems and burns. Airbags have clearly been a successful

measure to mitigate the severity of injuries in a frontal car accident, so much so that some vehicles are now fitted with nine separate airbag systems covering a range of accident types.

2:3.5.2 Highland and Islands Airports Limited Remote Airfield Light Switching for MEDEVAC[11]

The airbag comparison considers the risk versus benefit to the driver and passengers of a vehicle. It is only slightly less clear when the argument is balancing direct quantitative units against probabilistic likelihood of a less frequent outcome in the same units transferred to another group of people. The example given below for remote operation of aerodrome lighting for a specific activity balances the number of lives saved each year, against the probability that a trained crew will have a crash.

In remote areas of Scotland when a member of the public suffers a medical emergency, such as a stroke or heart attack requiring major medical facilities, they often have to travel great distances to obtain medical services. There is a network of small airports[12] operated by Highland and Islands Airports Limited to serve the communities but they do not operate 24 hours per day, 365 days per year. To enable access by MEDEVAC Flights at some locations there has been a remote light switching system installed to illuminate the airport and enable operation of the flights. This type of system has risks in that certain types of failure at key points could potentially cause an aircraft crash, leading to fatalities. The operation of this system saves several lives a year. The operation is approved on the basis that the larger number of lives saved, versus the probable number of lives lost by aircraft accident, means that on balance there is a net number of lives saved.

2:3.5.3 National Institute for Health and Care Excellence.

The National Institute for Health and Care Excellence[13] (NICE) has a difficult responsibility for assessing new medical treatments and recommending whether or not they should be widely available on the UK National Health Service.

NICE evaluates new medical treatments and assesses how much additional ‘quality life’ can be expected as a result of the treatment, termed a quality-adjusted life year (QALY). This measure combines both quantity (length) of life and health-related quality of life. This additional quality life benefit is determined by expert judgement based upon assessment of the medical evidence. This evidence may come from clinical trials and reports of the new of the treatment both elsewhere in the world and previous usage in the UK. Many new medical treatments have severe debilitating side-effects ranging from short-term sickness, skin sensitivity, hair loss, to longer term effects such as numbness in the nerves, impact on hearing, etc. The additional quality life includes the side-effects and is then considered against the cost of the treatment when compared to existing treatments to determine whether it should be made widely available. The costs are allocated into three bands where the lowest cost generally is recommended, the middle band is subject to further assessment and some treatments are recommended, and the top band where the cost per additional year of life is deemed too high.

This is effectively the inverse of a safety argument as the argument is that the likely life extension is worthwhile for the cost. There is a bounded set of costs per QALY which inform the acceptable cost per QALY for a given assessment.

[11] MEDEVAC stands for MEDical EVACuation, and is sometimes known as MEDIVAC (MEDIcal eVACuation)
[12] Barra, Benbecula, Campbeltown, Dundee, Islay, Inverness, Kirkwall, Stornoway, Sumburgh, Tiree and Wick
[13] https://www.nice.org.uk/

2:3.5.4 Individual Risk Versus Benefit Perceptions

It is important when considering the balance of risks that the goal is to achieve something, and any activity or inactivity has risks associated with it. When you undertake an action to reduce a health risk, such as sport to get healthy there is an increased risk of injury. This is illustrated by the UK House of Commons debate 6th April 1977 [21] on the increase in speed limits in the UK which restored the dual carriageway speed limit to 70 mph. The exchange summarises the need to balance safety with achieving our goals and that individual perception influences judgement.

Mr. John Ellis (MP for Brigg and Scunthorpe)

"Does my right hon. Friend agree that speed is a contributory factor to road accidents and that as a direct result of his announcement today more people will be killed and injured on our roads? Are we not all guilty of having a schizophrenic approach to this matter? It is no use the House or the public throwing up their hands in favour of safety when such an approach is adopted. Is it not only fair to say that?"

Mr. Rodgers (Secretary of State for Transport 1976-1979)

"Yes, it is only fair to say that. All of us, including me, suffer from schizophrenia. We want to save life, but we like driving fast. Although we should all travel slowly, with a red flag in front of us, people do not choose to do that. We must strike a balance. It is dangerous in some respects, but that is life."

This balance is in activities that we conduct on a daily basis, and each individual may have a different appetite for risk, complicating it. Decisions taken by public bodies are influenced by public opinion, which depends on a huge range of factors [22]. Where organisations make decisions on risk and how it is balanced then they need to do so in a structured manner.

2:3.5.5 Autonomous Vehicle Risk Distribution

Risk versus benefit questions when considering the introduction of autonomous vehicles (AV) focus on the benefit of increased driving capability. AVs have been positioned as improving transport accessibility for those who currently lack the capability or infrastructure to drive. Other improvements – such as environmental benefits from potential AV car-sharing schemes – have also been postulated, as have certain dis-benefits (e.g. the environmental implications of increased road traffic, and the subsequent effect on road network efficiency).

However, the risk associated with these benefits is more difficult to quantify. The technologies and real-world trials of AVs are not yet sufficiently advanced to allow a full understanding of the risk that they pose. While it is arguable that AVs will not gain societal acceptance until they demonstrably pose a lower overall risk than that associated with human drivers, this perspective masks differences in risk distribution. It is likely that the risk distribution associated with AVs will differ from that associated with human drivers, with AVs posing a disproportionate apportionment of risk to certain segments of society. For example, current sensor technology is less effective at detecting cyclists and pedestrians than a human driver, leading to fatal incidents [23].

In this example, a risk-benefit argument must take into account that the AV passenger has gained the benefit of increased transport capability, at the cost of an increased proportion of risk borne by non-drivers (cyclists, pedestrians). This is of particular interest to AV engineers and manufacturers, who must factor this into a decision about whether to prioritise the safety of an AV passenger or another road user in the event of a situation where the AV must choose which accident it selects where there is no option not to have an accident.

2:3.6 Framework of a Risk-Benefit Argument

A risk-benefit argument needs to be clear and transparent, as it should be challenged by the stakeholder accepting the risk, support presentation to any regulator and be the basis of any future defence should one be required. A risk-benefit argument could be structured as in the example below:-

1) Define the Risk:
   Describe the risk in terms of:
   a) Harmful Outcome – How hazards could lead to the outcomes.
   b) Severity – What is the consequence of the outcome, and the justification for that selection?
   c) Likelihood - What is the probability that the Harmful outcome will occur?
   d) Control Measures – What mitigations have been employed to prevent the outcome or reduce its severity?
   e) Duration – Whether Permanent or Temporary. If the risk is temporary what is the expected duration; what is expected to mitigate the risk in the future; and what is the confidence that the date will be achieved?
2) Explain practicability of further risk reduction:
   Confirm that redesign or additional mitigations cannot easily be employed to reduce the risk further. This should record the reasons why lowering the risk is not practicable (e.g. disproportionate cost / time / not possible) and what options were considered , (e.g. construction of new concrete bunker for a security risk, removal of external network connections for a cyber risk, operational limitations to constrain service delivery for a safety risk).
3) Expected Benefits:
   Describe the envisaged benefits.
4) Argue risk versus benefit:
   The section should clearly present an argument as to why the benefits justify the level of risk achieved. This should be in terms of:
   a) Description of comparison methodology for assessing the risks versus benefit.
   b) Justification of the comparison methodology (include weaknesses of selected method).
   c) Argument of the risk versus benefits balance.
      i. Include Assumptions.
      ii. Quantitative where meaningful.
      iii. Include counter-evidence.
   d) Where there are alternative options the argument should show why the selected option is the preferred solution. This might be supported by the use of a table showing the respective weightings of the options.
5) Monitoring Arrangements:.
   Describe how the increased level of risk is monitored, including the assumptions and any trigger points for escalation or action.
6) Conclusion:
   Summarise the risk versus benefit and present recommendation for their acceptance or rejection.

The risk-benefit argument is likely to be one of the last parts of an assurance case being produced prior to operational deployment of a change. It can be included in the assurance case or as a separate artefact in the overall suite. Once prepared it can be presented to the duty holder and, where required, the regulator.

2:4 MODULAR ASSURANCE

2:4.1 Introduction

2:4.1.1 Problem Statement

Large and complex assurance cases, which contain multiple assurance arguments, can be difficult to follow and comprehend. They can be unfocussed and a reviewer may find it difficult to determine whether all the necessary aspects have been considered and whether their consideration is sufficient within the context of use of the *[system]* that is the subject of the case. This is particularly apparent where multiple teams, within the same organisation or in external organisations, collaborate to generate an evolving assurance case. In this scenario, the relationships between independently developed aspects of the assurance argument may be left implicit and unstructured, leading to an incoherent overall argument where it is not easy to determine whether the interfaces between elements of the argument constitute complete coverage of the scope, whether duplication or omission has occurred or where there are mismatches or incompatibilities at the interface.

A consequence of assurance cases that are difficult to comprehend, is that it is very difficult to assess the impact of a change to the subject of that assurance case, on the argument itself. As a result a small change to a system, for example, could result in a significant effort to assess and/or address the impact of change to the assurance argument. Where rigorous arguments are necessary, this can introduce a cost penalty for small changes in that their assurance effort is disproportionately high; or alternatively, changes to the system may be inhibited by the disproportionate cost of updating the assurance case. Optimally, the impact on the assurance case should be in proportion to the size of the change to the system. In essence, modularity should help to make more transparent the impact of change to the overall argument. This facilitates options for reuse of system elements and also for dealing with related variant systems or services which represent a small delta from the originally assured system, product, service, etc. Use of evolutionary design and development processes, such as ‘agile’, with their cadence of change, exacerbates the above problems, as does increasing complexity and novelty of our systems. Hence, they all require a considered approach to use of assurance cases if the advantages are not to be offset by increased effort required for assurance of the system.

This paper aims to provide guidance on producing assurance arguments that demonstrate technical clarity and also clarity on ownership of specific elements of the argument.

2:4.1.2 Scope

This guidance paper deals with the structure and presentation of large and complex assurance arguments. It does not address the content or nature of the arguments that should be made, but rather how to present them in a way to support reader comprehension. This benefits in that the impact of change to a *[system]* might more easily be identified within the assurance argument, reducing the effort required to assure the modified *[system]*.

2:4.1.3 Structure

This paper sets out the rationale and purpose in use of modular assurance cases as well as the enhanced emphasis on interfaces (Section 2:4.2). Guidance is then provided on defining and optimising assurance case architectures and the complexities and pitfalls of composing arguments are discussed (Section 2:4.3 and 2:4.4). Certain properties that are challenging to argue in a modular way are discussed, as well as specific considerations for high assurance entities (Section 2:4.5 and 2:4.6). Finally, practical advice is offered in terms of how the assurance case is reported, so as to maintain the modular approach and comment is made on the availability of tool support for this approach (Section 2:4.7 and 2:4.8).

2:4.2 Good Practice - Concepts & Principles

This guidance paper recommends the use of modularity in assurance arguments as a means to improve comprehensibility of the argument and to support the identification of the impact of change following a modification to the subject of the assurance argument.

Modularity supports two key types of structuring, 'basic' and 'for compositional arguments', which may be used individually or in combination. Additional concerns and guidance for using modularity are provided.

2:4.2.1 Basic Structuring

Basic structuring allows the argument author to focus the reader's attention on the intended part of the argument structure, whilst handling potentially distracting 'side' arguments in a separate argument module. For example, one module of an argument might deal with the failure reporting mechanism for an in-service system, another may present an argument about the configuration management system. Where a non-trivial argument is necessary to support the use of a particular process cited within the main argument, for example, such as an argument about the suitability of a particular analysis technique, that could be distracting to the main argument. The overall argument would be read and understood more effectively by the reader if the 'side' argument were presented in a different module, perhaps one dedicated to all processes used in development of the system. This is commonly referred to as 'separation of concerns'.

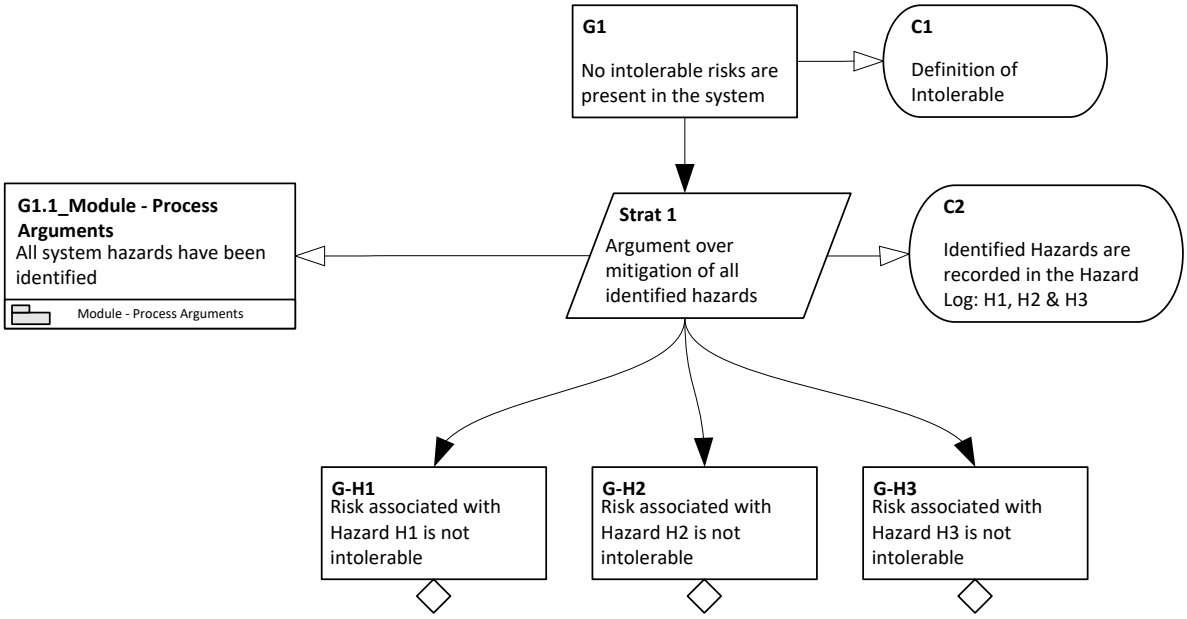


Figure 2:4-1 - Example of a 'side argument' expressed using GSN 'Away Goal'

A recommended separation of concerns might be 'risk', 'confidence' and 'conformance' as described in Section 2:5. In text-only assurance arguments, for example, this may be largely achieved by initially declaring the structure of the argument and then reflecting that in the structure of the report and through appropriate choice of section headers, annexes or related artefacts aligned to the topics of each module of argument.

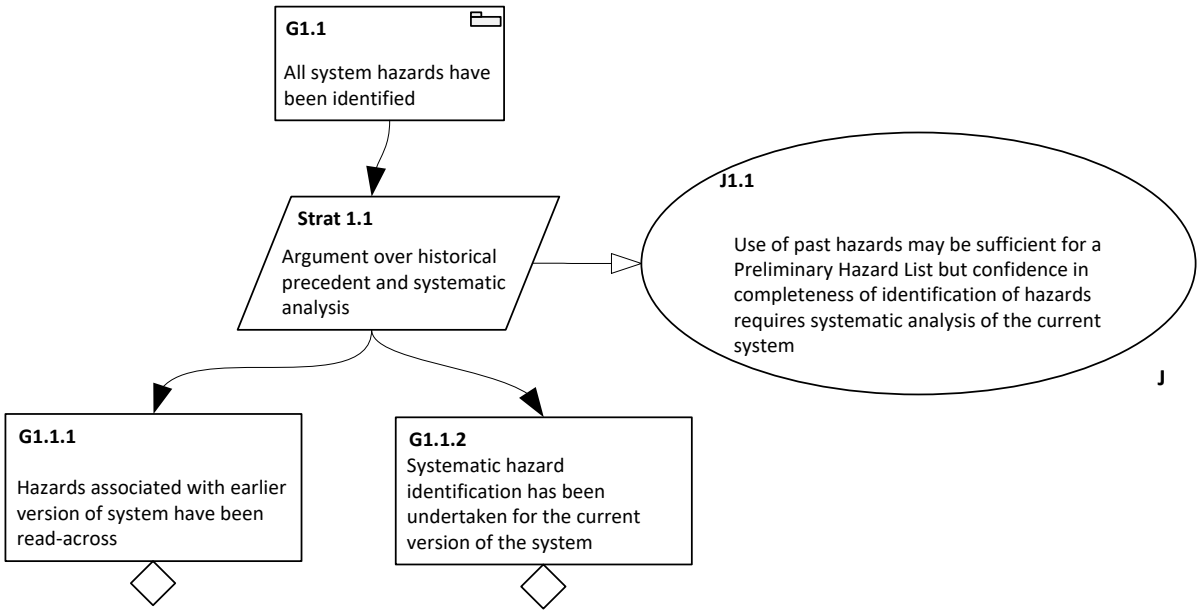


Figure 2:4-2 - Example of the 'side argument' expressed in a separate GSN module

2:4.2.2 Structuring for Compositional Arguments

Many systems are developed by integrating components, software modules/applications or similar, that are being/have been developed and produced by disparate teams, either within or external to the organisation that is developing the overall assurance argument. Similarly in a service context, for example, a safety-related aircraft maintenance service might include contracted-out services provided by specialist suppliers, such as ejection seat servicing, as part of the overall service for which an assurance argument may be needed. When an overall *[system]* is composed by integrating those separate parts, it would be advantageous if the assurance argument could be similarly composed from elements of assurance arguments produced for those parts. This approach also better supports the reuse of elements as the impact of the changed *[system]* can be more readily assessed and the requirement on replacement elements can be more readily understood.

2:4.2.3 Interfaces

A necessary consequence of 'modularising' an assurance argument is that interfaces in the argument will be created between the modules. These interfaces should record any needs, dependencies or shared assumptions between the linked modules. The assurance required from the overall argument will dictate the proportional response to the rigour of definition of the interface.

Interfaces need to cover the functional behaviour that is to be assured or depended upon between modules, and also other explicit or, often, implicit information. For example, assumptions about how a system will be operated may have a significant impact on how an argument is presented, but different assumptions might be made between parties working independently. A check of compatible assumptions and presumptions must be made at any interface when composing the overall argument. Note that identical interface assumptions may not be necessary, providing the combination of subordinate arguments do not reduce the scope of argument that is claimed for the overall case.

It is very easy to overlook quite simple interface compatibility issues, such as units of measure, endiansim, reference datum, standards compliance, etc. Being explicit about interfaces encourages checking of these issues.

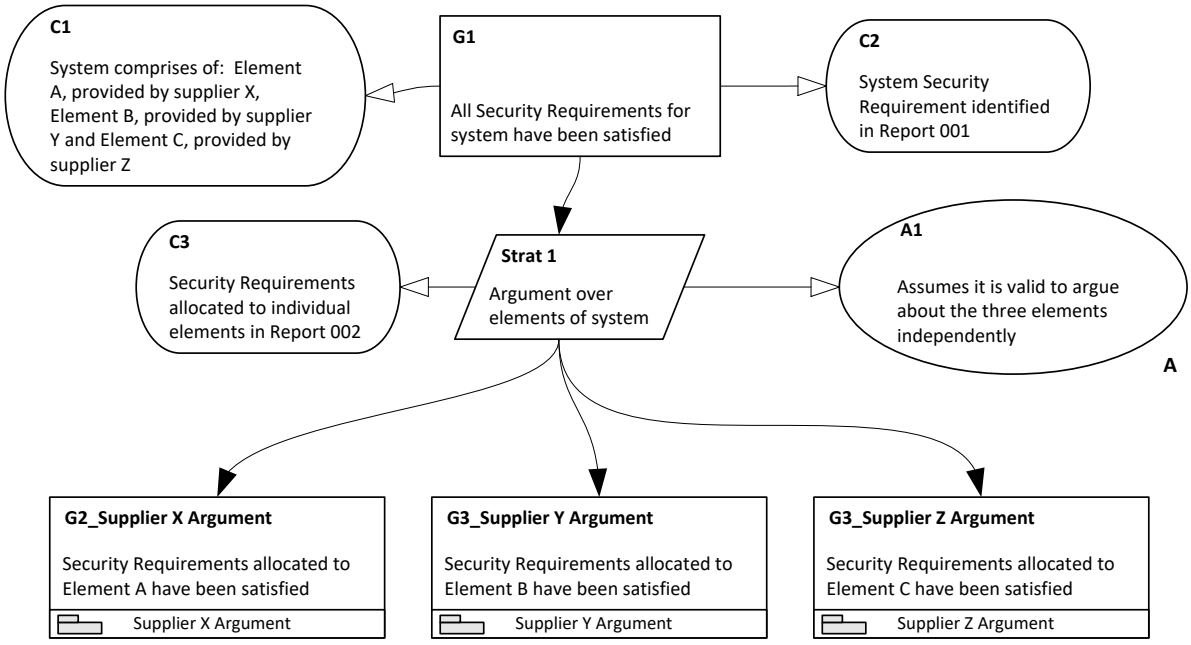


Figure 2:4-3 - Example of a GSN Argument Composed from Arguments Provided by Different Third Parties

2:4.3 Guidance

Use of modularisation appears to be a simple and logical decision; however, there is a number of areas that strongly impact on the effectiveness of the approach. These are addressed in the following sections.

2:4.4 Structuring/Architecting the Argument

As with all aspects of assurance arguments, the level of effort expended in determining and optimising the argument structure should be proportionate[14], not only to the level of risk presented by the *[system]*, but also, in this instance, to the potential opportunity for re-use. The remainder of this section assumes a moderate level of assurance is required for a *[system]* which is non-trivial and for which there is expected to be opportunities for reuse.

2:4.4.1 Defining and Optimising the Architecture of the Argument

The structure of an assurance argument should be considered early in the design lifecycle, and, for complex entities, is ideally considered as part of the selection criteria for deciding upon the *[system]*'s architecture, such that mutual optimisation between design architecture and assurance argument architecture can take place. For example, a system or software architecture which makes it difficult to delineate between software provided by two different third parties would make it difficult to allocate requirements to each of those supplier's software and hence difficult for them to contribute independently to the overall assurance argument.

A range of techniques may be used to optimise a *[system]*'s architecture, such as the SEI Architecture Trade-Off Analysis Method [24]. This method identifies a number of candidate system architectures and evaluates them against a set of criteria that is identified for the individual project. One of those criteria is likely to include an evaluation of the impact of change on the architecture by considering a number of change scenarios. It is feasible to extend this method to propose a number of assurance argument architectures per proposed *[system]*'s architecture and extend the evaluation criteria to consider the impact of change on the assurance arguments.

[14] A paper on this topic is planned for a future update of this guidance document.

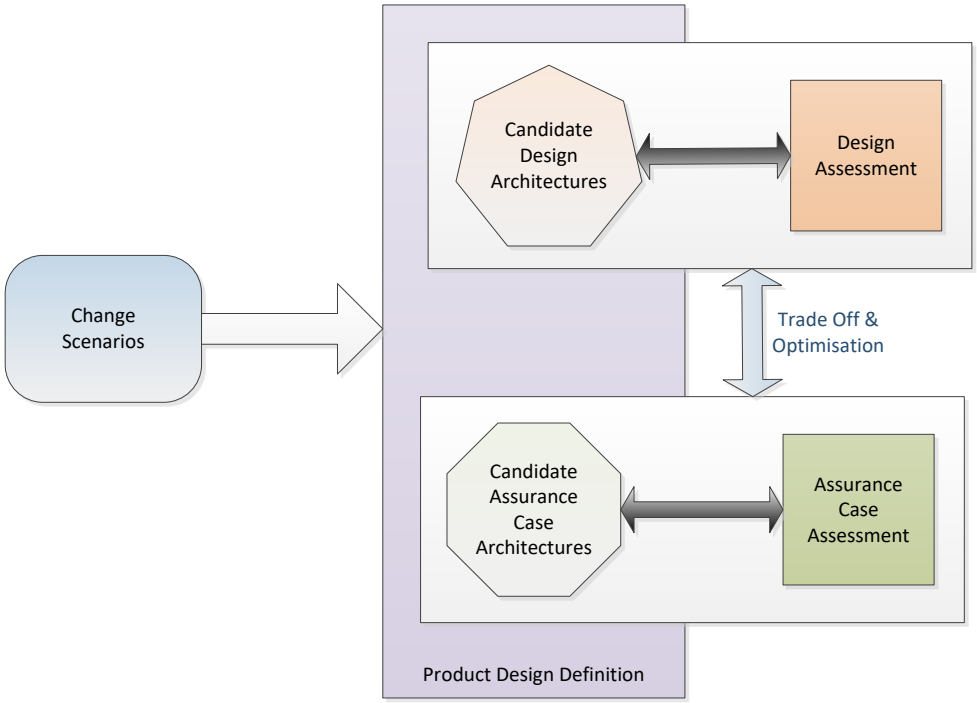


Figure 2:4-4 - Concept of Mutual Trade-Off and Optimisation Between Design Architecture and Assurance Architecture Choices

Note that it is feasible for this type of analysis to recommend changes to design architecture to facilitate an assurance argument architecture that deals more appropriately with change. Note also that the assurance argument does not necessarily have to follow the same architecture as the design. It may be more appropriate, for example, to make an overarching argument about a development process rather than making an argument about process for each of the developed elements of a *[system]*'s design. Part of the trade-off consideration should be associated with the cost to define the interfaces needed to support modular assurance, which will be higher for higher assurance entities. For example, more assurance case modules and interfaces in high change areas bring benefit of change containment, even in high assurance entities, whereas there is little value in defining interfaces between low assurance elements of design with low anticipated change where modularity will bring minimal benefit.

2:4.4.2 Abstraction and Hierarchy in the Assurance Argument

The architecting of an assurance argument has been compared to the architecting of a system, and similarly, other systems and software engineering considerations might apply, such as abstraction and information hiding. The ideas of 'separation of concerns' described earlier could also be achieved by a hierarchical structure which may hide information that is not required by the flow of the current argument. Whilst this is easier to visualise in a graphical notation, the same could be achieved in a text only argument. For example, the interface of a module of text argument may be indicated as presenting an argument about the process used to generate evidence for a system. The actual argument within the module may talk about how the process was selected, reviewed and approved and perhaps the competency of the people using the process, but those aspects were chosen to be hidden at the interface to simplify readability of the main argument.

2:4.5 Considerations for Integration of Modular Assurance Cases

There are specific considerations that may become necessary to facilitate integration of assurance case modules and challenges that arise from trying to partition system level properties using a modular approach.

2:4.5.1 Challenges of Composing Arguments

When creating an argument composed from other modules of argument, developed in isolation, it is necessary to consider whether the product, in operation, also operates in an isolated way. If, say, the behaviour of one component can interfere with the operation of another component, it may not be valid to simply present the assurance arguments made independently. For example, if one component can utilise resource that would then be unavailable to another component, that is a form of interference that would compromise the associated assurance argument. As well as being clear on resource requirements at the interface, in this example, a high level argument and evidence would typically be required about how resource allocation in the system would be policed and how each component is intended to behave when it has insufficient resource available. Other common non-interference arguments that may need to be made, include interference from modules of different safety assurance or security accreditation levels. This might be addressed by partitioning arguments from the operating system, for example.

For complex entities, it may even be advantageous to include intermediate 'integration' arguments for a sub-set of related argument modules to improve clarity and minimise mutual reliance between modules, leading to better change isolation. These will typically not map to a physical/logical design element and are purely an artefact in the assurance case domain.

2:4.5.2 Properties that are Challenging to Handle in a Modular Way

Certain properties can be difficult to address in a modular way; end-to-end performance measures being a typical example. Timing properties could be handled by budgeting for each individual component of a software system, for example, but is likely to yield a pessimistic outcome. Probabilistic timing analysis may be sufficient for low assurance systems or higher assurance systems with soft deadlines, but will be challenging where hard timing deadlines exist. This needs to be considered as early as possible in the design lifecycle. Overall computing resource may also be difficult to handle in a modular way, as an additional example. Again, a simple summation of all resource demands that could be placed is likely to yield a pessimistic requirement for resource, so approaches like advanced scheduling techniques may be necessary to deploy, but their assurance needs to be argued at a higher level than the independent components.

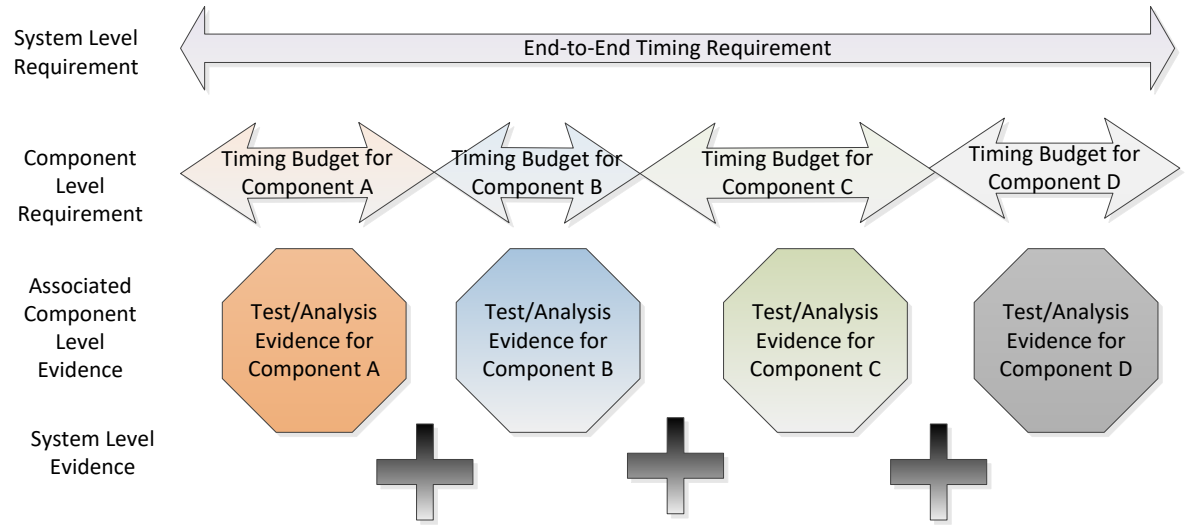


Figure 2:4-5 - Providing Assurance Evidence for Timing Behaviour by Allocating Budgets to Each Contributing Component

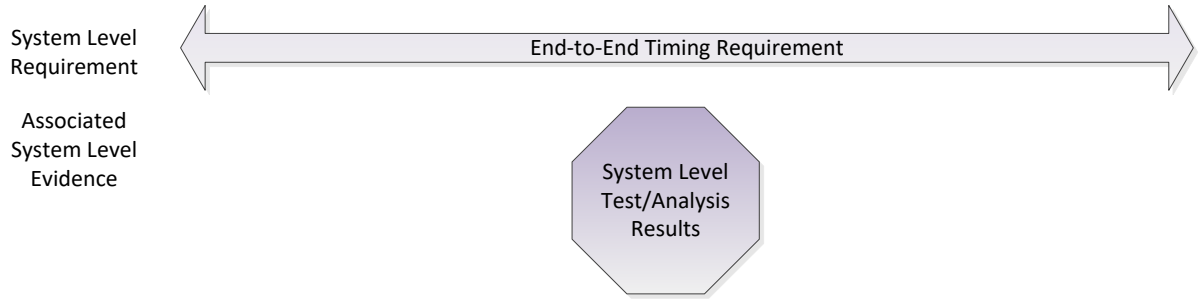


Figure 2:4-6 - Providing Assurance Evidence for Timing Behaviour by Only Analysing/Testing at System Level, (Potentially 'Breaking' Modularity)

2:4.6 Further considerations for High Assurance entities

In high assurance entities, the rigour with which an interface needs to be defined will require significant effort. This effort needs to be factored into the method used for deciding the assurance argument architecture that is optimal to present the argument for a specific design architecture. The need for each interface in the argument should be critically considered as part of the optimisation process to determine whether it adds flexibility or incurs unnecessary effort/cost to produce.

2:4.7 Structuring the Modular Assurance Argument Report

Consideration of how the assurance argument will be finally reported is a further consideration. To support reusability and change management, it may be desirable to report on each module of the argument in a separate, stand-alone artefact. Traditional configuration management techniques would suggest that all artefacts that cooperate to provide an overall argument mutually reference each other, at a specific issue version. This would result in a situation where, if one artefact issue version is changed, it will necessitate the revision of all of the other linked artefacts, which may entirely circumvent the intent to reduce the impact of change. In such cases, consider a single, simple, 'indexing' artefact[15], which records the collective issue status of all linked artefacts, but which will have to be changed every time any of the artefacts change and do not reference out to any other artefact in the individual assurance argument module reports.

2:4.8 Tool Support for Modular Assurance Arguments

The concepts described above can be exercised within a text based assurance argument which is supported only by a text editing tool, however, a number of notations exist which provide advantages for presenting an assurance argument. 'Goal Structuring Notation' [5] and 'Claims, Argument, Evidence', [6] are key examples, both of which can be mapped to the Structured Assurance Case Metamodel specification, (SACM) [7]. Tool support for producing large and complex structures provides some degree of consistency checking and traceability, depending on the tool, which can be highly advantageous. Unfortunately, at the time of writing, very few tools fully support modularity in assurance arguments but it is hoped this will be incorporated in due course and also with the integrity to support use with the highest assurance arguments.

[15] Similar in concept to a Master Record Index used in the Aerospace sector.

2:4.9 Further Information

The 'Industrial Avionics Working Group' (IAWG) has developed an example modular assurance case process and a suite of guidance including 'Modular Software Safety Case Process Description, MSSC 201' [25].

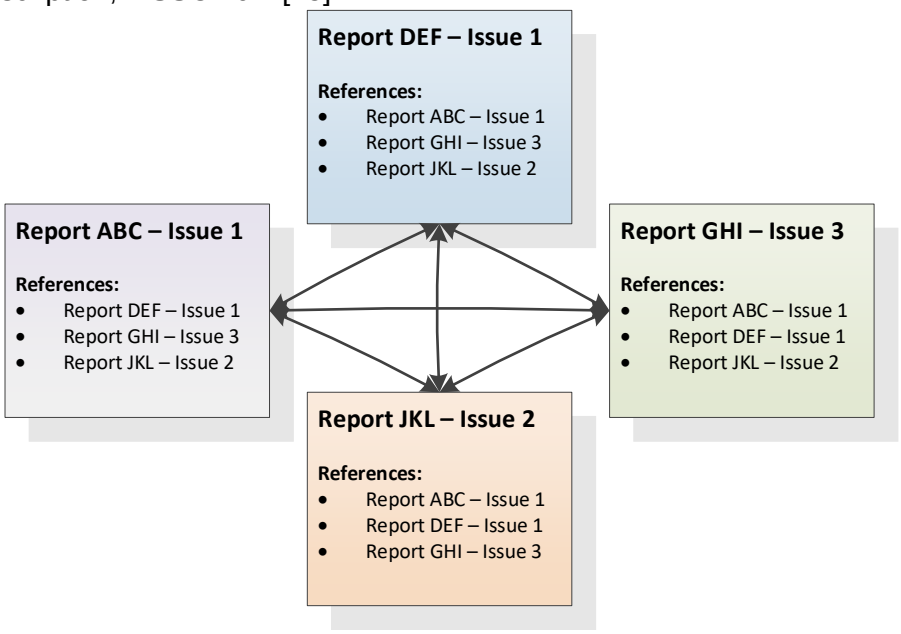


Figure 2:4-7 - Showing the Interdependencies Between Artefacts Created by Cross-referencing

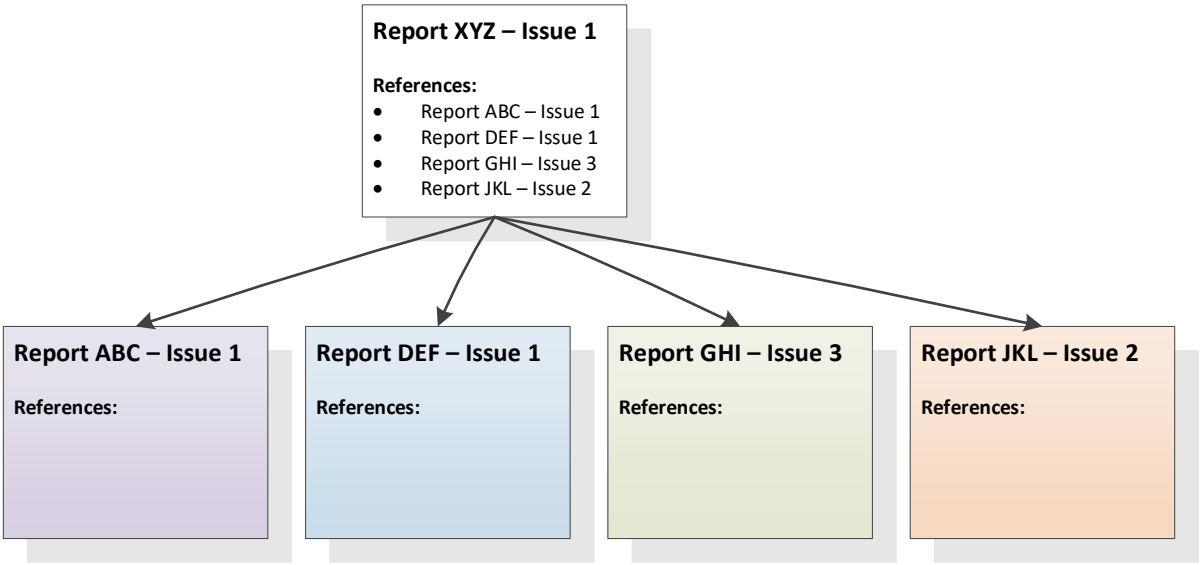


Figure 2:4-8 - Showing an Alternative Approach where Referencing Happens in Only One Place

2:5 RISK-CONFIDENCE-CONFORMANCE APPROACH

2:5.1 Introduction

2:5.1.1 Problem Statement

A compelling assurance case will require an argument and evidence covering many different aspects of the assurance of the *[system]* and property being assured. Present practice is generally to develop a single, unified assurance argument that does not distinguish between arguments of risk, confidence or conformance. This practice merges what are essentially three different but interrelated arguments. These elements are essential to a compelling assurance argument, but presenting them all in an intermingled fashion typically results in a larger (often rambling) argument and makes grasping the crucial structures difficult for the reader. Clarity of presentation is important for all stakeholders even though their interests might differ. For developers, the distinction between the risk, confidence and conformance arguments will help provide clearer direction on the steps involved in constructing each argument and a better understanding of the necessary development and assurance steps. For reviewers, the distinction will help focus attention on those aspects of the argument that are weakly supported.

This paper is developed from ideas published by Hawkins et.al. [26]

2:5.1.2 Structure

This paper first defines in Section 2:5.2.1 three different types of argument that may be present within an assurance case. Section 2:5.3 provides guidance on how arguments of each type may be created. Section 2:5.4 provides more detailed discussion on the way in which a confidence argument may be structured, and Section 2:5.5 provides a realistic example of risk and confidence arguments for a hypothetical insulin pump.

2:5.2 Good Practice

An assurance case should include three separate, but inter-connected arguments. These are:

1. A ***risk argument*** that records the arguments and evidence used to establish direct claims of the acceptability of risk
2. A ***confidence argument*** that justifies the sufficiency of confidence in the risk argument
3. A conformance ***argument*** that justifies belief in conformance with the requirements of a standard or regulation

Figure 2:5-1 illustrates the relationships that exist between these three types of argument.

To ensure a clear separation between the three types of argument while ensuring coherence of the assurance case as a whole, a number of simple criteria have been established for distinguishing between the different argument types. These criteria are detailed below.

2:5.2.1 Types of Argument

When developing an assurance case, the following criteria should be used to establish risk, confidence and conformance arguments. Further guidance on developing each type of argument is provided in Section 2:5.3.

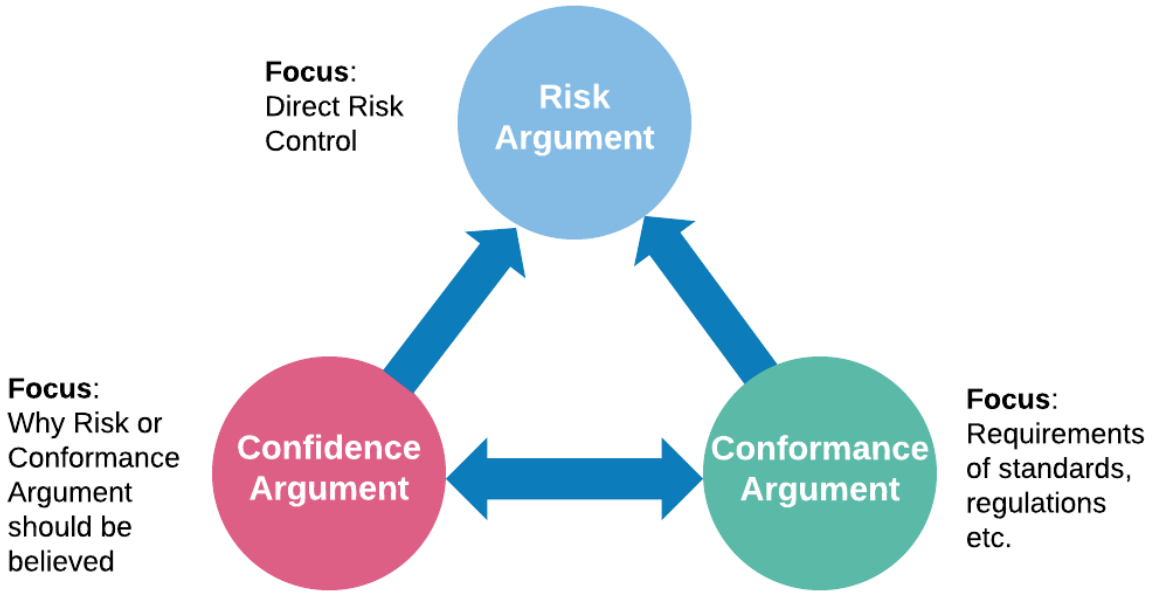


Figure 2:5-1 - Types of assurance case argument

2:5.2.1.1 Risk Argument

A risk argument records the asserted arguments and evidence of risk reduction. A risk argument should meet the following criteria:

- Everything cited in the risk argument should have a direct role as part of the causal chain to a risk.
- All claims in the risk argument should be claims about the *[system]* or its elements, properties, or properties of elements thereof.
- Artefacts from development (e.g. test reports and, by extension, their contents) should be referenced in the risk argument only as evidence or context.

2:5.2.1.2 Confidence Argument

A confidence argument records the reasons for having confidence in the risk argument. A confidence argument should meet the following criteria:

- Confidence argument claims should only address elements within the risk or conformance argument.
- Confidence argument claims may not be required for every assertion in a risk argument.

2:5.2.1.3 Conformance Argument

A conformance argument records the reasons for believing the requirements of a particular standard or regulation have been met. A conformance argument should meet the following criteria:

- Everything cited in the conformance argument should have a direct role in demonstrating that the requirements of a standard, regulation or other similar artefacts have been met.
- The conformance argument should not make direct claims regarding the sufficiency of risk reduction (which should be provided in the risk argument).
- Conformance arguments should reference elements of the risk or confidence arguments where these have a role in demonstrating conformance.

2:5.3 Guidance

2:5.3.1 The Need for Multiple Arguments in an Assurance Case

The present practice of including in a single argument elements that record direct arguments of risk mitigation, supporting arguments that are 'confidence-raising' and arguments regarding conformance with standards leads to a number of difficulties including:

- Arguments tend to become large and unwieldy, because there is too much information that is irrelevant or whose role is unclear. The entry criterion for the inclusion of an argument or item of evidence in the assurance argument is often, 'Does this have any possible bearing on the assurance of the *[system]*?'. All three types of argument are admitted by this criterion. This can lead to voluminous, rambling, ad infinitum arguments.
- All three types of argument tend to be poorly prepared, because the lack of distinction between them makes it more difficult to spot incompleteness or poor structure in any of them.
- Necessary elements of the argument are sometimes omitted, because the need for the specific elements is lost in the volume of the argument.
- Arguments become indirect and unfocused, and the link between elements of the argument and risk is often lost.
- Unnecessary material is sometimes included in arguments without proper consideration or explanation of its relevance – 'just in case'.
- Arguments become difficult to build, and weaknesses of the argument are sometimes not evident and so are easily overlooked.
- Arguments become difficult to review because of the size and lack of focus.

These difficulties are serious since they all detract from the basic purposes of using assurance cases. Separation of the arguments offers the opportunity to mitigate these difficulties by providing different foci for risk, confidence and conformance. In addition, careful attention to linking the arguments provides a mechanism for guiding analysis of the interrelationship between risk, confidence and conformance.

2:5.3.2 Risk Arguments

An assurance argument must explain how the available evidence supports the overall claim of acceptable assurance. Best practice, risk-based assurance arguments decompose this claim into arguments that justify the acceptability of the identified risks. The argument states what 'adequately' addressed means for each risk and then identifies the evidence supporting the conclusion. This structure explains the purpose of each piece of evidence. A truly risk-based assurance argument must always be focused upon the identification and mitigation of risks in this way. For a safety case, where the risks being considered are those associated with hazards, everything cited in the risk argument should therefore have a direct role as part of the causal chain to the hazard. That is, all of the goals in the assurance argument must be claims about the *[system]* or its elements, properties, or properties of elements thereof.

Artefacts from development (e.g. test reports and, by extension, their contents) may be referenced only in solution or context elements. Strict adherence to this tight definition of a risk argument ensures the focus of the assurance case is clearly on the direct management of risk.

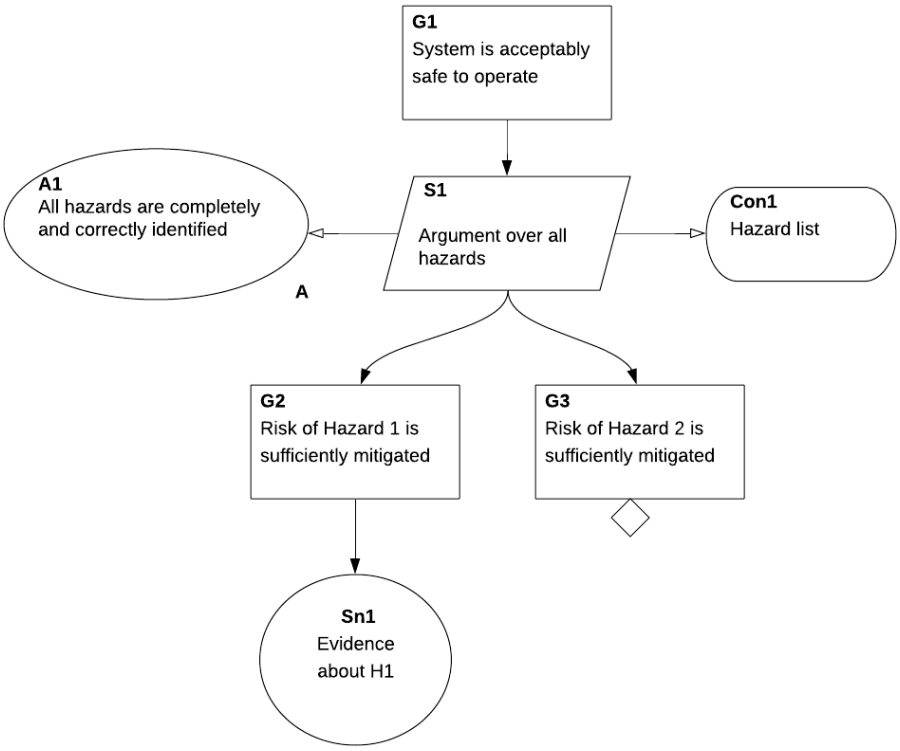


Figure 2:5-2 - An example risk argument represented using GSN

Figure 2:5-2 shows a simplified example of a risk argument represented in GSN, in which a claim is provided for acceptable safety of the system. Although a risk argument should always have mitigation of risk at its core as shown in this example, in practice arguing the acceptability of overall risk is more complex than simply arguing the acceptability of each individual identified risk.

2:5.3.3 Confidence Arguments

Both evidence and argument will typically be imperfect. Having sufficient confidence in the claims made in the risk argument is therefore essential. A confidence argument demonstrates the justification for confidence in a risk argument. There will be uncertainties associated with aspects of the risk argument or supporting evidence. The role of the confidence argument is to address those uncertainties explicitly and explain why there is sufficient confidence in the risk argument. Any argument includes a number of assertions. These assertions relate to the sufficiency and appropriateness of the inferences declared in the argument, the context and assumptions used, and the evidence cited.

A recorded risk argument is merely a recorded position that collects together these assertions. To be compelling, the confidence argument must justify the truth of the assertions made. If an argument assertion cannot be justified, then the argument will not be believed (it will not provide the required assurance). The confidence argument provides the justification for argument assertions. In order to indicate the assertion in the risk argument that the confidence argument is associated with, the confidence argument may be tied to a number of Assurance Claim Points (ACP).

When using GSN, for example, an ACP can be indicated with a named black rectangle on the relevant link[16]. A confidence argument is developed for each ACP. It is important to note that an explicit confidence argument may not be required for every assertion in a risk argument. It is important instead to identify those assertions for which it is either felt to be most important to demonstrate confidence (for example because they play a critical role in the risk argument or are unusual assertions), or those assertions for which confidence is more difficult to

[16] The representation of ACPs in GSN is included for information.

demonstrate. Those assertions for which a confidence argument is not supplied do not require an ACP.

Figure 2:5-3 shows the risk argument from Figure 2:5-2 with the addition of some ACPs named ACP1, ACP2 and ACP3. These ACPs correspond to three different types of assertion:

- asserted inference (ACP1)
- asserted context (ACP2)
- asserted solution (ACP3).

Note that in Figure 2:5-3, an assumption (A1) about the completeness and correctness of the identified hazards is no longer required due to the provision of a confidence argument at ACP2. Below we discuss each of these three types of assertion in more detail.

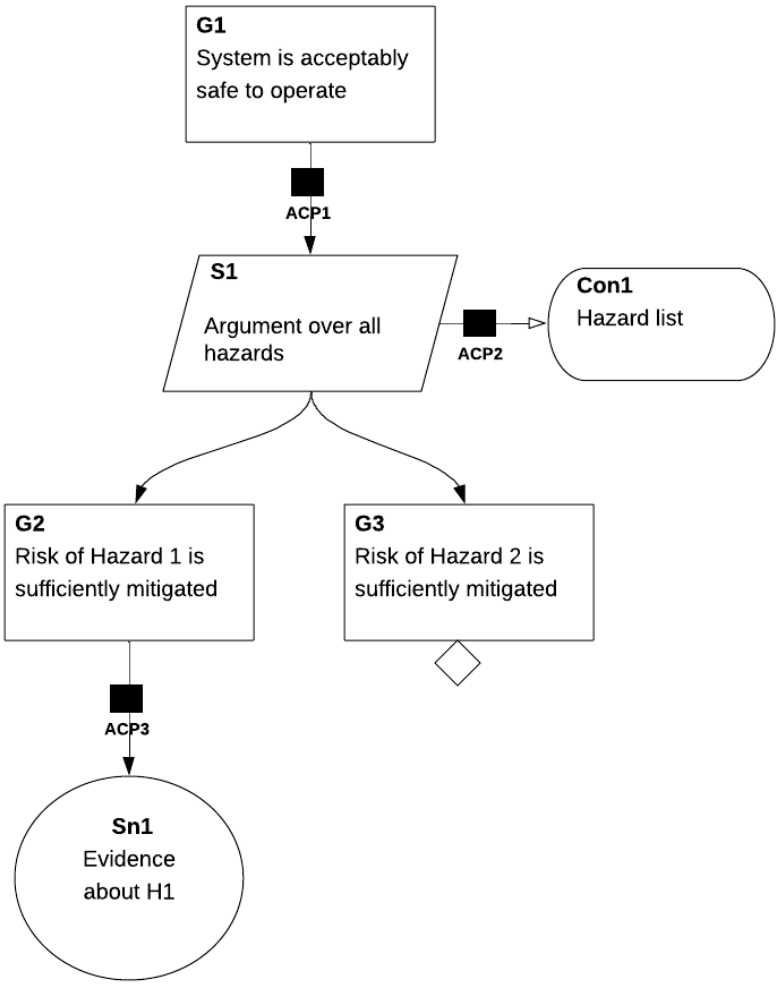


Figure 2:5-3 - Examples of the use of ACPs in GSN

2:5.3.3.1 Asserted Inference

Each time a claim is said to be supported by other claims in an argument, an assertion is being made that the inference is appropriate and sufficient. Only in deductive arguments do premise claims prove a particular conclusion. Instead, for inductive arguments, the assertion is that the probable truth of the premises is sufficient to establish the probable truth of the conclusion.

Although assurance cases can contain a mix of both deductive and inductive arguments, inductive arguments typically pervade. For example, Figure 2:5-4 shows (in GSN) the assertion that, given the applicable context, the sub-claims put forward to implement the chosen argument strategy are, if true, a sufficient basis upon which to infer the conclusion stated in the parent claim.

To gain assurance in the adopted argument strategy, it is necessary to provide a confidence argument that demonstrates why the asserted inference should be believed. The ACP for an asserted inference is the link between the parent claim and its strategy or sub-claims.

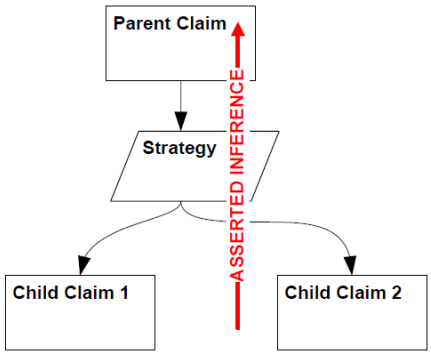


Figure 2:5-4 - Asserted inference

For example in Figure 2:5-3, the asserted inference is that if the risk of all hazards are sufficiently mitigated then the system is acceptably safe to operate. The role of the confidence argument for ACP1 is to demonstrate why it should be believed that the two supporting claims of hazard mitigation are sufficient to draw the overall conclusion about system safety. Details on how such a confidence argument may be constructed are provided in Section 2:5.4.

2:5.3.3.2 Asserted Context

Each time contextual information (represented by context or assumption elements) is introduced into the argument, it is being asserted that the context is appropriate for the argument elements to which it applies. For example, consider a context reference to a list of failure modes for a particular piece of equipment. The introduction of this context element when arguing about the assurance of that piece of equipment implicitly asserts that the list of failure modes referred to is appropriate to the application and operating context in question.

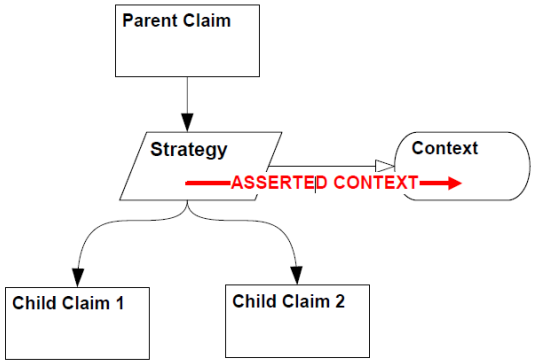


Figure 2:5-5 - Asserted context

Figure 2:5-5 shows asserted context for an argument strategy. The assurance of the strategy depends upon the confidence that the context or assumption stated is appropriate for that strategy and its sub-goals. It is necessary to provide a confidence argument that demonstrates why it should be believed that the asserted context is appropriate. In the example in Figure 2:5-6 it is being asserted that the hazards given in the referenced hazard list are the relevant hazards. For this context to be appropriate there must be confidence that the hazard list is appropriate with respect to the system, application and context. The role of the confidence argument at ACP2 is therefore to demonstrate why it should be believed that citing this hazard list defines the appropriate context at this point in the risk argument. The ACP for asserted

context is the link to the contextual element. Details on how such a confidence argument may be constructed are provided in Section 2:5.4.

In addition, since context provides a reference to an artefact, it may also be necessary as part of the confidence argument to provide an argument as to the trustworthiness (or integrity) of the artefact to which the context refers. In the example in Figure 2:5-3 this would provide an argument and evidence as to why there is confidence that the hazard list does not contain errors or omissions. This confidence argument must therefore consider aspects such as the sufficiency of the hazard identification process that has been followed. The ACP for the trustworthiness of the artefact may be applied to an element that references the artefact, as shown for ACP4 in Figure 2:5-6. Details on how such a confidence argument may be constructed are provided in Section 2:5.4.

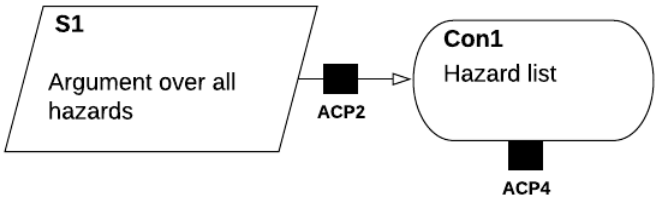


Figure 2:5-6 - Example of an ACP for a referenced contextual artefact

2:5.3.3.3 Asserted Solution

Each time evidence is referenced as a solution to the argument, it is being asserted that the evidence put forward is sufficient to support the claim. Figure 2:5-7 shows an asserted solution to a risk claim. The assurance of the solution depends upon the confidence that the evidence is appropriate to support the claim. The ACP for asserted solutions is the link to the solution element.

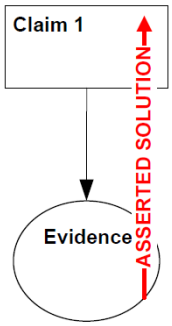


Figure 2:5-7 - Asserted solution

In the example shown in Figure 2:5-8 it is being asserted that the stress testing results are sufficient to demonstrate that the defined operational forces can be tolerated. For this solution to be sufficient there must be confidence that the stress testing performed is good enough for this purpose. The role of the confidence argument at ACP3 is to provide this confidence. This will involve considering whether the stress testing of the type being referred to is adequate to support the claim being made, as well as whether the particular tests performed were appropriate. Details on how such a confidence argument may be constructed are provided in Section 2:5.4.

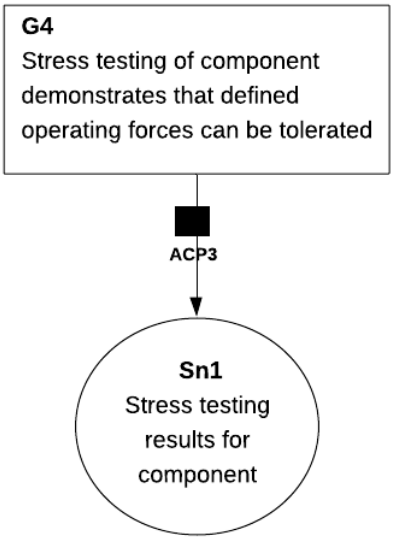


Figure 2:5-8 - ACP relating to an asserted solution

In addition, as for asserted context, solutions provide references to artefacts, so it may also be necessary as part of the confidence argument to provide an argument as to the trustworthiness of the artefact to which the solution refers. In the example in Figure 2:5-9 this would provide an argument to justify the trustworthiness of the stress testing results, such as the rigour with which they were performed and the competence of the people performing the tests.

The ACP for the trustworthiness of the artefact may be applied to a solution that references the artefact, as shown for ACP5 in Figure 2:5-9. Details on how such a confidence argument may be constructed are provided in Section 2:5.4.

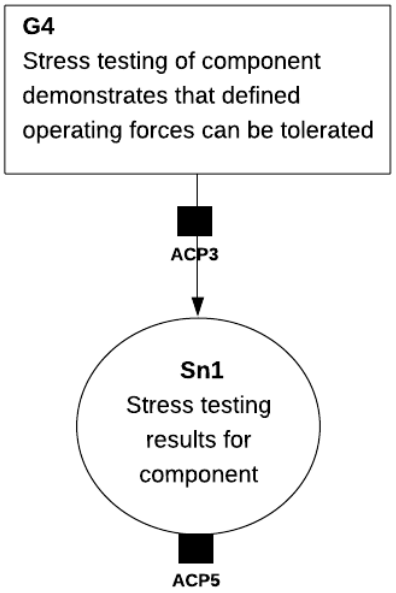


Figure 2:5-9 - Example of an ACP for a referenced solution artefact

2:5.3.4 The Overall Confidence Argument

The individual fragments of confidence argument, each addressing a particular ACP in the risk argument, should be assembled together to form a single overall confidence argument (to

accompany the single risk argument). In addition to this simple structure, there is a number of potentially important concerns at the level of the overall confidence argument.

Firstly, arguing the sufficiency of the overall confidence in the risk argument can be more complex than the simple composition of arguments of sufficient confidence for each argument assertion (in the same way that arguing the acceptability of overall risk is more complex than simply arguing the acceptability of the risk posed by each individual hazard). For example, shortfalls in confidence in one part of the risk argument may be compensated by other arguments and evidence elsewhere in the argument. This must be addressed at the level of the overall confidence argument.

Secondly, it is useful to examine and justify whether the multiple lines of argument offered in the risk argument (undesirably) share common underlying uncertainties (i.e. there are common modes of failure in the argument).

Thirdly, for large risk arguments it may simply not be practical to provide arguments of confidence for every assertion in the risk argument. Instead, some selection and prioritisation of the assertions of the risk arguments to be covered by the confidence argument may need to be performed. This prioritisation would be done most appropriately by addressing those assertions relating to the most significant arguments of risk reduction in the primary risk argument. Care must be taken when making any decisions regarding parts of the confidence argument to omit.

2:5.3.5 Conformance Arguments

Although demonstrating conformance with standards, regulations or other similar artefacts is never, in itself, sufficient to form a compelling assurance argument for a *[system]*, it is often a necessary aspect of an assurance case. Demonstrating conformance to standards is often not straightforward. Problems stem mainly from the need to interpret the text of a standard to fit the specifics of a particular application. Although recording conformance is commonplace, the quality, transparency, and scrutability of artefacts can vary significantly. Using explicit, rigorous, and structured conformance arguments helps to explain how conformance has been demonstrated, as well as clarifying the relationship to the overall risk argument.

Standards will often contain requirements that must be interpreted in the context of each *[system]*. There are at least four distinct scenarios in which interpretation is necessary:

- the use of high-level goals in the standard;
- deliberate non-specificity in the standard including optionality and tailorability;
- the possibility of meeting the letter but not the intent of the standard;
- using a standard outside of its intended context (such as a different domain).

Neither self-assessment nor independent assessment is always straightforward and unambiguous in such cases. Compelling claims of conformance may therefore require greater exposition and transparency than existing conformance practices provide. Making a conformance argument explicit enables careful review – by developers, regulators, third parties, or some combination of these – which can find defects in its reasoning. Because these defects might hide instances of non- conformance, confidence in conformance can be raised by detecting and repairing them.

A conformance argument justifies belief in conformance, even if there's no compelling reason to believe that conformance is adequate evidence of assurance. As a result, the first level of decomposition in a conformance argument is over the standard's requirements (as opposed to direct claims of risk reduction). Claims that each requirement has been satisfied are further decomposed until each sub-claim can be supported by evidence. The chain of reasoning in a conformance argument both illustrates the developers' interpretation of the standard and defines what each item of evidence must show if a specific *[system]* is to conform to a given standard.

Due to the wide variation in the nature of requirements in standards, conformance arguments can be related to specific features of the risk mitigation argument, for example the achievement of a specified risk target, or a requirement to implement particular risk mitigation features. Conformance arguments may also concern matters of confidence, for example mandating the use of certain techniques or processes according to integrity level. In such cases the relevant argument to be used to demonstrate conformance may already exist in the risk or confidence arguments respectively. It would not be expected that these arguments would be duplicated in the conformance argument, but rather that the conformance argument would make reference to existing arguments in the other parts of the assurance case.

2:5.3.5.1 Example Conformance Argument

This section presents a simplified extract of a conformance argument relating to the DO-178B standard [27]. Figure 2:5-10 shows how software for a commercial aircraft meets one of the objectives of DO-178B.

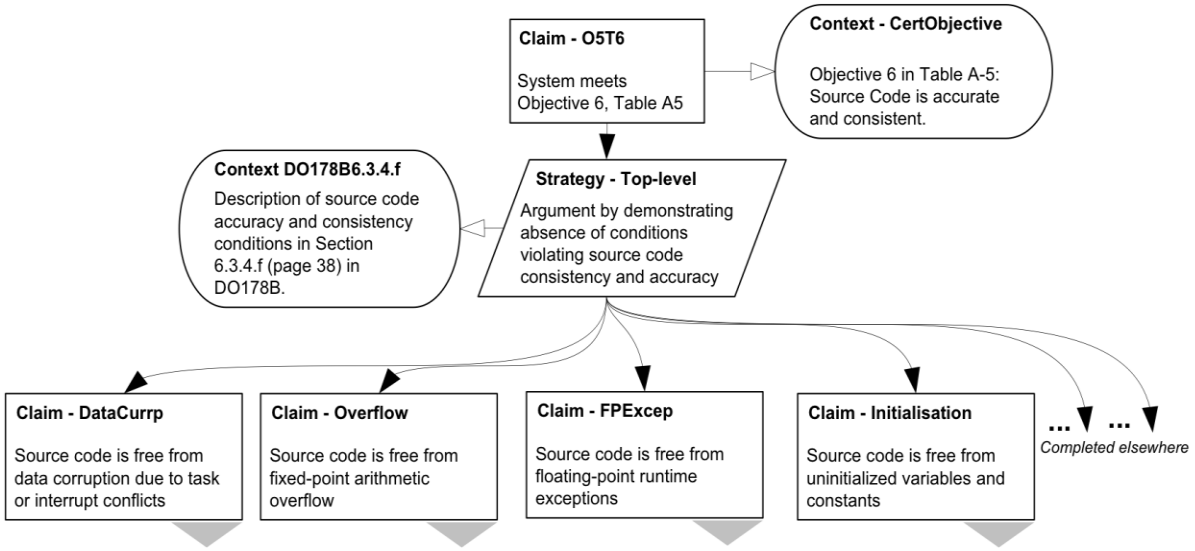


Figure 2:5-10 - A conformance argument to demonstrate conformance to an objective of DO-178B

Figure 2:5-11 presents the rationale for believing that the requirement that floating-point exceptions won't be raised at runtime is satisfied.

2:5.4 Structuring Confidence Arguments

2:5.4.1 Confidence Arguments Relating to Assertions

There are many ways in which a compelling confidence argument relating to an asserted inference in a risk or conformance argument could be made. The detailed nature of the argument will depend upon the *[system]* for which the risk argument is made and the specific nature of the assertion being considered. In general however, the confidence argument would often be expected to consider the following three areas:

1. There are grounds to support the probable truth of the assertion.
2. Residual uncertainties (assurance deficits) in the assertion have been identified.
3. The residual uncertainties (assurance deficits) in the assertion have been mitigated such that they are insufficient to cause concern.

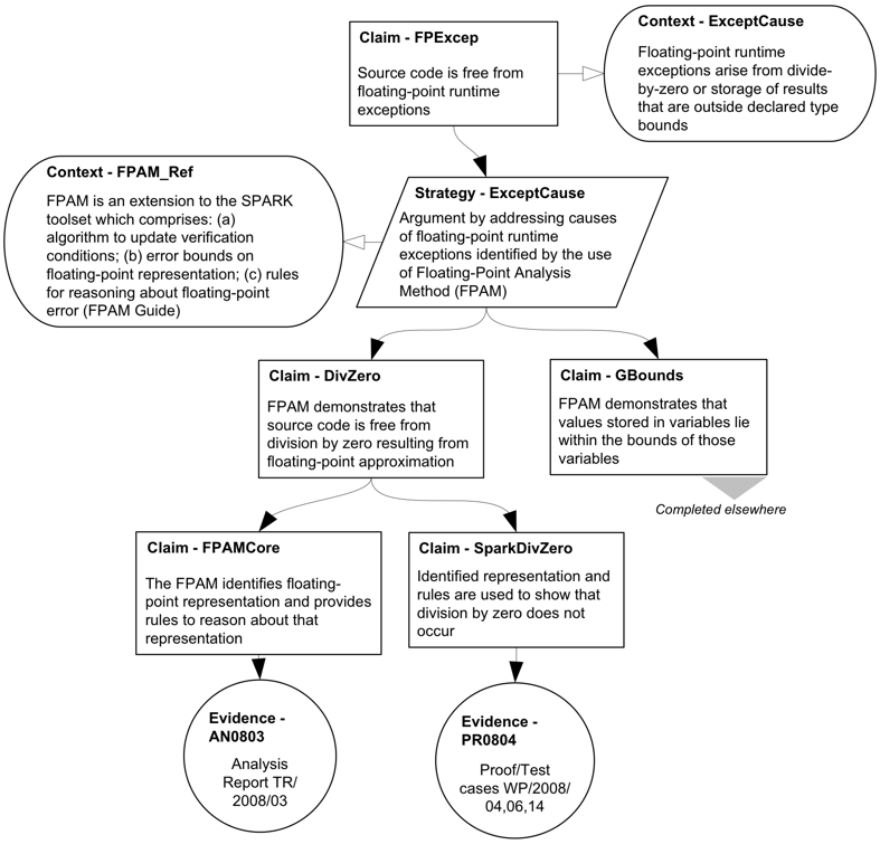


Figure 2:5-11 - An argument demonstrating how a requirement of the standard is satisfied

Consideration area 1 considers the reasons why the assertion should be believed. This aspect is realised as the decomposition of a goal of the form 'the assertion <x> is true'. As in any argument, goal decomposition continues until the goal can be supported with evidence. Unlike the risk argument, however, the goals in this portion of the confidence argument are typically expected to be claims about properties of development artefacts (i.e. 'process' claims). For example, the decomposition of a solution assertion goal might contain arguments over the properties of test plans, development tools, and configuration management systems. Goal decomposition in this portion of the confidence argument should continue until no reasonable observer would deny that the artefact cited offers positive evidence in support of the claim.

Consideration area 2 involves justifying that the uncertainties (assurance deficits) surrounding the assertion have been identified. The identification of an assurance deficit identifies a gap in our knowledge relating to an assertion in the argument. One reason that assurance deficits are of interest is that they represent 'blind spots' in the argument, i.e. areas of the argument where no evidence has been presented. Should these 'blind spots' be eliminated (by providing the appropriate evidence) we may find that the evidence is positive (and supports the assertion made in the risk argument). However we may also find that the evidence is negative and forms counter-evidence to the risk argument. Recognising assurance deficits, therefore, helps identify the possible areas in the argument where counter-evidence may exist. (This guiding of the otherwise boundless search for counter-evidence is a useful side-effect of the identification of assurance deficits.) For example, consider a case where there is no control flow analysis evidence of the absence of infinite loops in some source code. When arguing that a return value will always be provided, we should consider the probability of the existence of counter-evidence to our claim (i.e. if we were to provide the control flow analysis – how probable is it that an infinite loop will be detected?).

Consideration area 3 requires that the identified uncertainties (assurance deficits) are mitigated such that any remaining uncertainties can be argued to be acceptable. It is possible to mitigate any identified assurance deficits by taking one or more of four actions:

1. making changes to the design of the *[system]*, e.g. adding a hardware backup when it is impractical to demonstrate with adequate confidence that software has the properties necessary to ensure system assurance;
2. making changes to operation of the *[system]*, e.g. by limiting the conditions under which the system is used;
3. making changes to the assurance argument, e.g. adding an independent source of evidence;
4. generating additional evidence for the confidence argument, e.g. increasing the coverage of software functional tests.

It is important to note that completely mitigating all assurance deficits is not normally achievable. In many cases it would be possible to go on forever generating additional evidence to try to gain some additional confidence. It is therefore necessary to make a judgment on when assurance deficits can be tolerated, and to justify this within the confidence argument.

Based on the discussion above, Figure 2:5-10 shows an example argument pattern that could be used for structuring a confidence argument for an asserted inference (e.g. ACP1 in Figure 2:5-3). An example of how this pattern may be instantiated is included in Section 2:5.5.

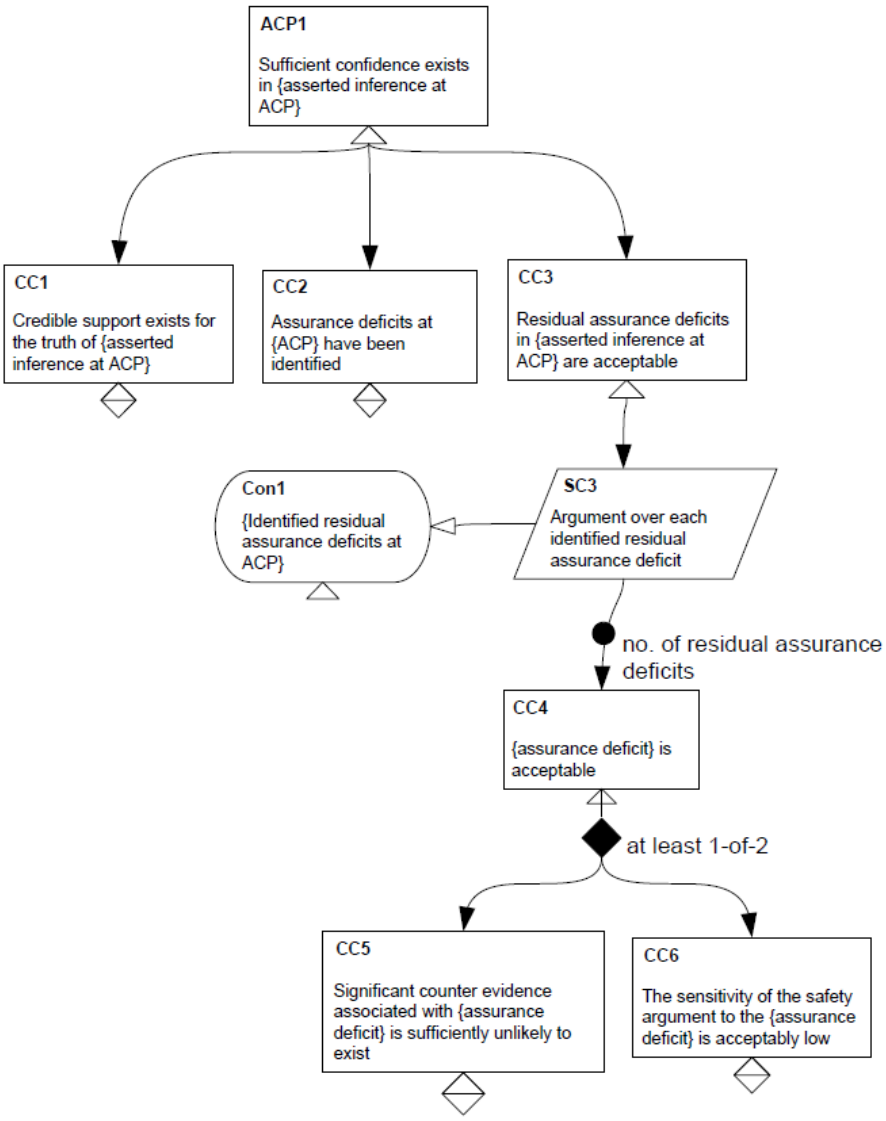


Figure 2:5-12 - Argument pattern for a confidence argument for an asserted inference

2:5.4.2 Confidence Arguments Relating to Referenced Artefacts

When providing a confidence argument for elements that make reference to artefacts (such as asserted solutions and asserted context), the confidence argument must also consider whether the asserted artefact is itself trustworthy. For instance, in the example shown in Figure 2:5-8, the trustworthiness of the stress testing itself must be considered in the confidence argument, e.g. how rigorous were the tests?, who performed them?, did they follow correct procedures? etc. This is in addition to the appropriateness and adequacy of stress testing in demonstrating that operating forces can be tolerated that was considered in the confidence argument for the asserted relationship. The confidence argument for the referenced artefact will follow essentially the same form as that for assertions. Figure 2:5-13 shows an example argument pattern that could be used for structuring a confidence argument for a referenced artefact. The argument under CC12 would follow the same pattern as CC3 in Figure 2:5-12.

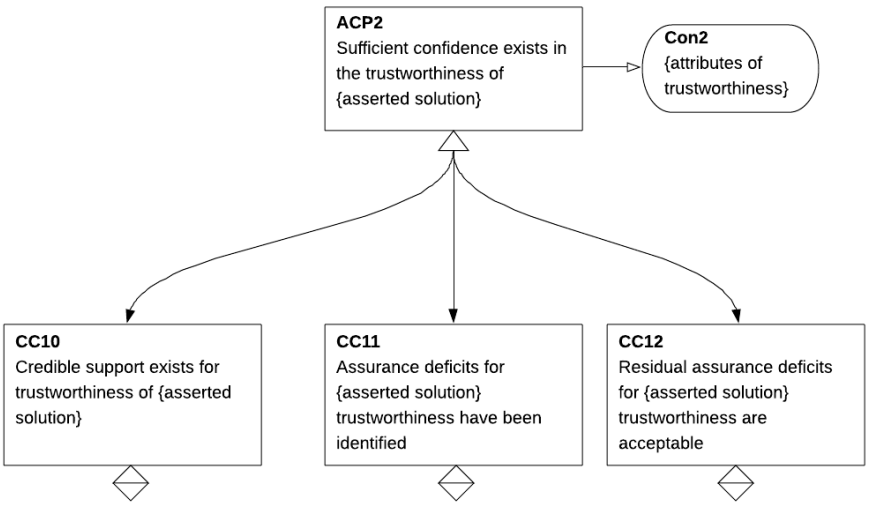


Figure 2:5-13 - Argument pattern for a confidence argument for a referenced artefact

2:5.5 Example Risk and Confidence Arguments for a Hypothetical Insulin Pump

In this section we illustrate the application of risk and confidence arguments through consideration of the safety case for a hypothetical insulin pump. The example has been simplified for clarity and is provided for illustrative purposes only. Figure 2:5-14 shows the high-level structure of the risk argument represented in GSN. The claim that the insulin pump is adequately safe for routine use is supported by arguing over each of the identified credible hazards to which the patient might be subject.

A number of ACPs requiring a confidence argument are identified:

- ACP1. There is sufficient confidence that the definitions of adequately safe and routine use are appropriate for the safety claim being made. If the scope defined by this context is not appropriate for the way in which the system is operated, for example if the device is used in an unplanned manner in a hospital, then the argument presented may not be valid.
- ACP2. There is sufficient confidence that the details of diabetic patient types and usage environments are appropriate for the intended use. Usage outside of the expected set of environments might invalidate the safety claim.
- ACP3. There is sufficient confidence that the pump design referred to is a sufficient representation of the actual pump used.
- ACP4. There is sufficient confidence that the pump design is accurately recorded.
- AC5. There is sufficient confidence that mitigating credible hazards will demonstrate that the insulin pump is adequately safe for routine use. Arguing over hazards is a

widely accepted strategy in safety engineering, and this fragment of the confidence case is simple to construct.

- ACP6. There is sufficient confidence that the correct hazard list is referred to and is up to date.
- ACP7. There is sufficient confidence that the list of credible hazards is complete and correct for the defined system. Inadequate definition of a hazard or omission of a hazard might invalidate the safety claim.

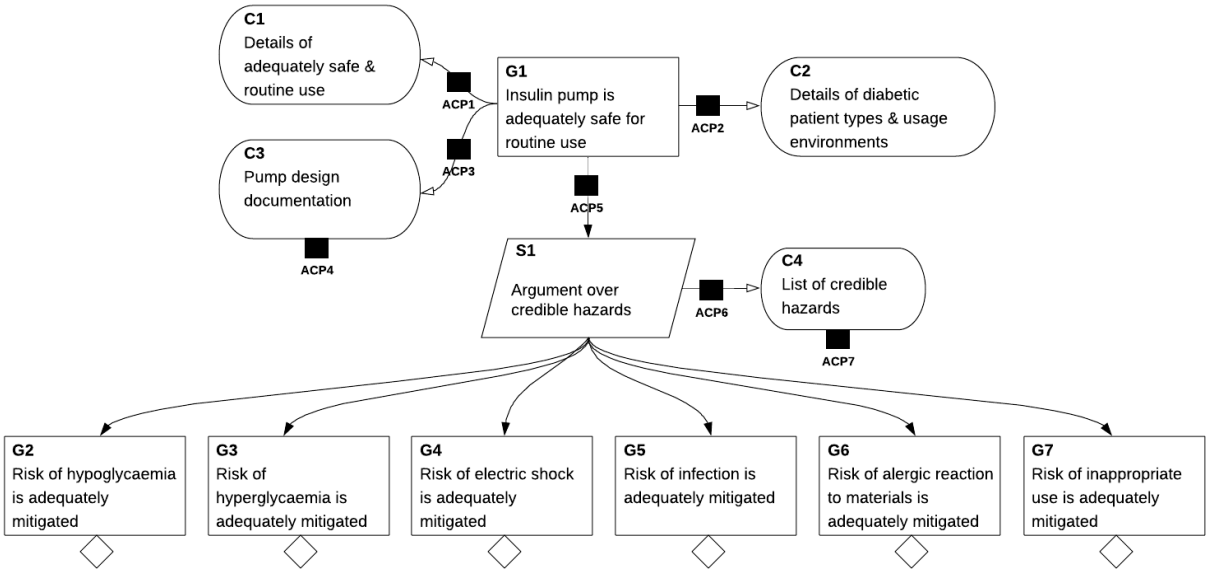


Figure 2:5-14 - High-level risk argument for an insulin pump

Here we illustrate ACP3 and ACP4 in detail. ACP3 relates to the appropriateness of the asserted context, so to create a suitable confidence argument we can use the pattern shown in Figure 2:5-12. ACP 4 relates to the trustworthiness of the artefact referred to by the context, so to create a suitable confidence argument we can use the pattern shown in Figure 2:5-13. Figure 2:5-15 shows the fragments of these arguments that deal with the assurance deficits for the appropriateness and trustworthiness of the pump design. The remainder of the pattern would be instantiated in a similar manner.

Figure 2:5-15 (a) shows a fragment of the confidence argument for ACP3. Here the assurance deficits that we associate with the appropriateness of the pump design documentation need to be enumerated and argument and evidence as to their acceptability provided. In this example, we consider just two assurance deficits:

- The possible deficit relating to the use of the correct design document (CC4), i.e., are we sufficiently confident that the correct document was actually referenced? Here this is mitigated through evidence of an effective configuration control process being used.
- The possible deficit relating to local modification of the pump by the user (CC7), i.e., if the user of the pump makes changes to it after it is put into operation then the pump will no longer correspond to the design document. Here this is mitigated by the fact that we have confidence that the users of the pumps (the patients) are trustworthy, and there have been no previous reports of users making unapproved modifications to the pump. In addition, the sensitivity of the assurance case to this deficit is low because of the built-in consistency checks of the pump itself that would be likely to identify any tampering by the user.

Figure 2:5-15 (b) shows a fragment of the confidence argument for ACP4. Here the assurance deficits that we associate with the trustworthiness of the pump design documentation need to

be enumerated and argument and evidence as to their acceptability provided. In this example, we consider just two assurance deficits:

- The possible deficit relating to the relative inexperience of the engineer responsible for creating the design document (CC13), i.e. are we sufficiently confident that they are competent enough that they won't introduce significant errors into the design. Here this is mitigated by ensuring that additional independent review of the design document is in place.
- The possible deficit introduced by the use of a commercial modelling tool (CC14), i.e. are we sufficiently confident that the document was not corrupted in some way by the tool itself? Here this is mitigated by the fact that there have been no reports of any significant deficiencies in the tool during previous usage, and by ensuring that manual inspection is also carried out of the design models produced in order to check for errors.

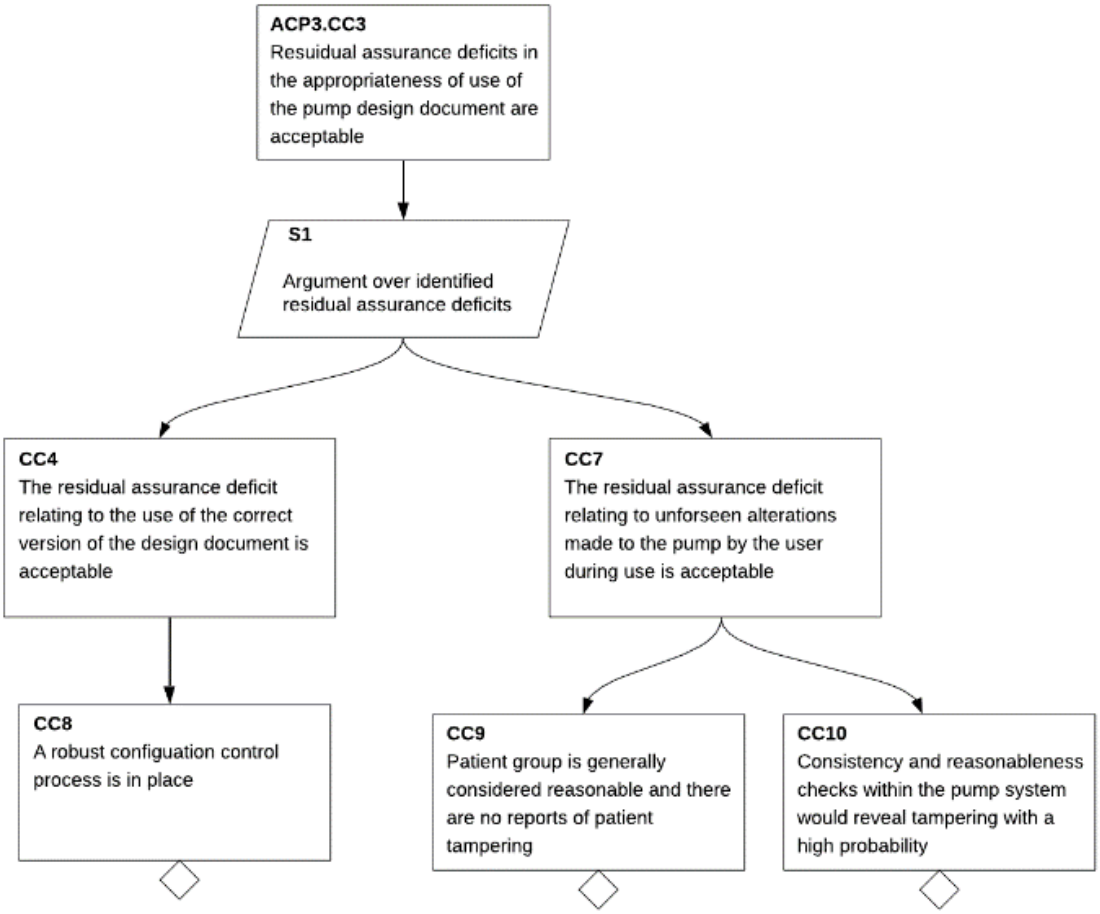


Figure 2:5-15 (a) - Part of the confidence arguments for ACP3

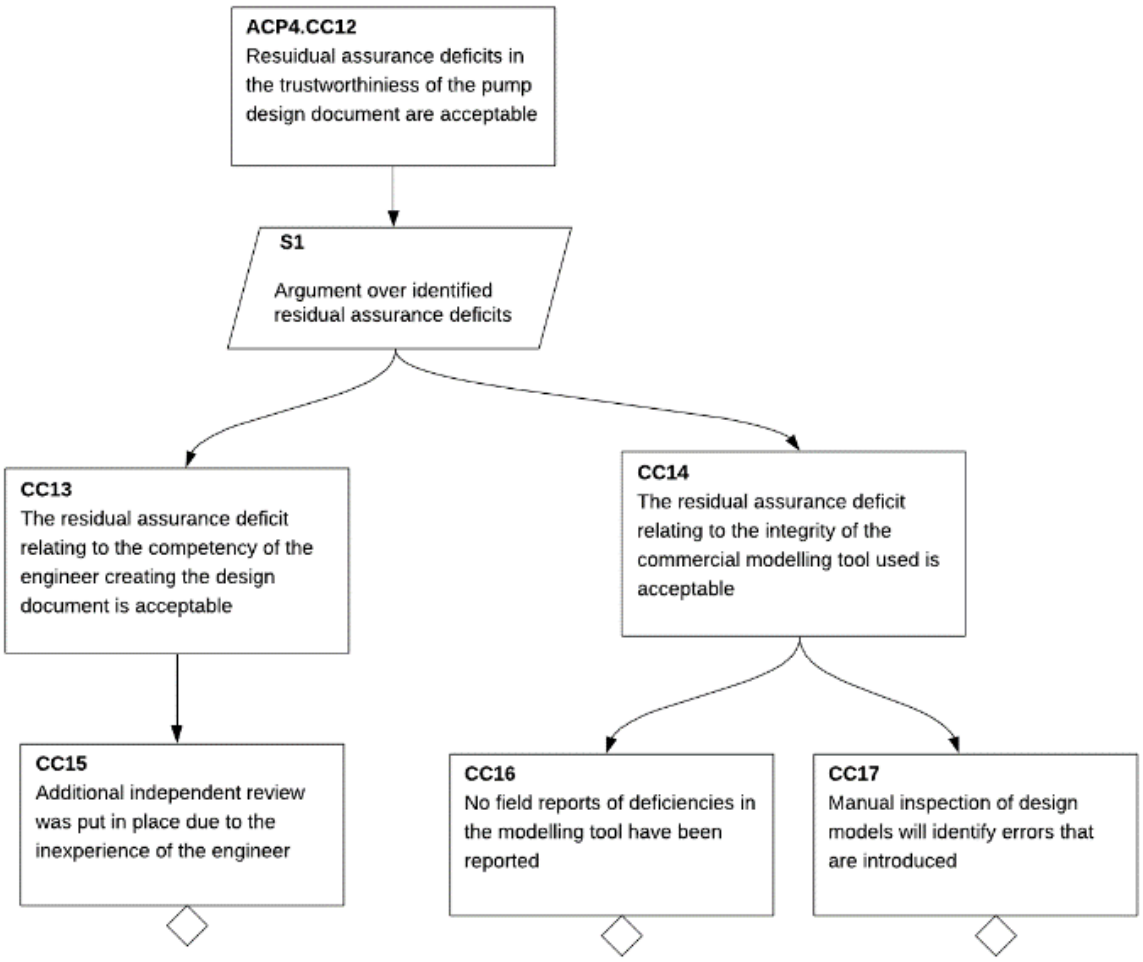


Figure 2:5-15 (b) Part of the confidence arguments for ACP4

2:6 DIALECTIC ARGUMENT

2:6.1 Introduction

2:6.1.1 Problem Statement

In the past investigation of the truth of claims made in an assurance argument has often been limited to recording and analysing of evidence that challenges some or all of those claims.

A number of issues, such as awareness and availability of methods and/or tools, have limited the adoption of a wider approach of challenging arguments as part of their creation; often such challenge is limited to the identification of counter-evidence e.g. arising from faults or incidents. While treatment of counter-evidence is important, it is a subset of the wider need for challenge.

This lack of adoption, together with emerging (and hence relatively untested) technologies, such as model-based development and autonomy, has the potential to reduce confidence due to the lack of robust challenge applied to ever more complex assurance cases.

The use of dialectics in assurance cases is relatively immature and lacking in public domain examples of real world use. The guidance presented in this paper builds on “Six Honest Serving Men for Today (Dialectic Argumentation for Assurance Cases)” [30], which raises the profile of dialectic argumentation, draws on existing dialectic discussion and suggests ways forward with techniques. [30] includes a literature survey that identifies prior material investigating the use of dialectic thinking for review of assurance arguments.

2:6.1.2 Scope

The guidance contained herein primarily seeks to introduce the concept of dialectic argumentation and show how it can be used to challenge an assurance argument while it is still being authored, as opposed to problems being found at a later stage by other reviewers, or worse, by an event that brings its trustworthiness into question. The benefit of use for review is also recognised.

2:6.1.3 Structure

After this introduction, this guidance:

- defines dialectic argument and demonstrates its purpose (Section 2:6.2.1);
- provides rationale and justification for the use of dialectic argument (Section 2:6.2.2);
- suggests what is good dialectic practice, by addressing:
  - Principles of use (Section 2:6.3.1).
  - Common terminology (Section 2:6.3.3).
  - Applications and examples (Section 2:6.3.4 and 2:6.3.5).
- Additional recommendations (Section 2:6.3.6).
- draws conclusions (Section 2:6.4);
- signposts useful background material on the topic of dialectic argument (Section 2:6.5).

2:6.2 Definition and Rationale

2:6.2.1 Definition of Dialectics

Dialectic argument is, in its simplest form, the **investigation of truth**. The root of the term “dialectic” is “dialogue”, implying that a dialectic process is two-way. From the Oxford English Dictionary [28], the definition of ‘dialectic’ is:

“critical investigation of truth through reasoned argument, often specifically by means of dialogue or discussion. (Noun)”

Applying this to the use of dialectic argument we can see that this is far more than just the consideration of counter-evidence, a common misconception regarding the use of the term. In fact, a 'dialectical model' can be created based on a set of 'moves' or responses to any part of the argument.

Applied to the assertions made within structured arguments, as used by modern assurance cases, dialectics compares options, tests truth by discussion, logically disputes and constructively criticises; the use of dialectic argument gives a framework for creating and challenging assurance cases.

2:6.2.2 Rationale for Use

Dialectic argument could be seen as an overhead bringing additional work, so it is helpful to outline some of the benefits by listing some weaknesses of existing assurance cases. These weaknesses are repeated at section 2:6.4, with how dialectics helps each one.

Assurance cases are by their nature subjective and ultimately require stakeholder acceptance of the subjective position. As a result there is a number of common criticisms that may be levied at an assurance case:

- Confidence in the conclusions of the assurance case may be lost due to facts which weaken the argument, including where evidence appears to support the argument but is in fact not trustworthy.
- Logical fallacies i.e. the assurance case may give the (deceptive) appearance of being a good argument without actually being so.
- There can be insufficient premise in support of the conclusion.
- Confirmation bias i.e. there can be a tendency to bolster a hypothesis by seeking consistent evidence while disregarding inconsistent evidence. The preference for hypothesis-consistent information can lead to a false outcome.

When constructing an assurance case, over-emphasis on the high-level claims can lead to bias and concentration of effort on demonstration of the assertions that directly follow. This can often result in insufficient focus on the wider picture. Whilst there is no deliberate intention to mislead, the outcome can inadvertently do so and hence bring about the above.

Explicit dialectic argument can redress the balance. Timely application can add strength to arguments that might otherwise go unchallenged until it is too late.

2:6.3 Good Practice

This guidance paper seeks to explain the concept, show where dialectic argument is useful, and advise on how to construct such arguments.

It must be recognised that argument authors have traditionally challenged what they write as they compose, just as any author would. What is proposed by this guidance is making such self-questioning more explicit and formal, with the added benefit of providing an audit-trail that can be used effectively as evidence of decisions taken along the way.

2:6.3.1 Principles of Dialectic Arguments

When constructing dialectic arguments, the following principles apply:

- Often the initial argument is a starting point for discussion; it should of course represent the author's best knowledge at the time, but is open to, and expected to be subject to, constructive challenge from other parties.
- The author should, wherever possible, try to pre-empt the return (dialectic) argument, and be explicit about where parts of the assurance case, including evidence or strategy, are insufficient or weak.

2:6.3.2 Application

Initially in producing an assurance case, a dialectic approach should be used by assurance case practitioners and authors. Once complete, an assurance case is challenged by review. Such reviewers, for example an Independent Safety Auditor/Assessor (ISA), the customers, the users, the regulators, and even an investigator could create their own dialectic argument to support and record the review process in a formal way.

The processes used for accident investigation sometimes use a dialectic-type approach, challenging previously undisputed or overstated assumptions e.g. around operator training and/or processes, and examining supporting evidence for weaknesses or absence. Applying such a method 'before the fact' to the construction of assurance artefacts, rather than after an incident can improve the likelihood of pre-empting unwanted outcomes.

For service applications requiring 'continuous assurance', their frequent incremental change and evolution brings the need for regression assurance to assure each change. Such assurance, or the justification for where it is needed, can be weak and so a dialectic approach may bring additional benefit to strengthen such assurance.

It may be difficult to produce a dialectic argument, when working in isolation. It is not always straightforward to appraise critically one's own original thoughts. A two-way dialectic communication may be beneficial in identifying the required underlying material. This may take the form of a "return argument", constructed separately to the main argument, or it may simply be the process of refinement applied to the main argument only. In practice, it is likely to be used differently according to those involved, and what works best for the task at hand.

2:6.3.3 Common Terminology

In other dialectic-related guidance and/or discussion (see Further Information in Section 2:6.5), specific terms are used to refine the type of challenge being applied. The main terms are as follows, noting that the exact definition and application of the terms is sometimes different and therefore is the subject of further clarification.

A dialectic argument is based on a process that applies challenges to any part of the original argument, and responses to those challenges, as a set of 'moves' or 'responses'. Challenges to the argument can be directed at any part of an argument. A challenge has the potential to defeat an argument and is referred to in some papers as a 'defeater', although it may not necessarily result in defeat.

Challenges may take the form of a rebuttal or may undercut. Challenges are applied in the following way:

- Rebuttal: by evidencing an argument which results in a conclusion counter to that of the assurance case;
- Undercutting: through the introduction of additional facts (evidence) that challenge the reasoning within the assurance case.

In some of the existing work a further form of challenge is defined:

- Undermining: specifically raising doubts about the trustworthiness of evidence.

2:6.3.4 Examples of Dialectic Argument

This example considers a typical assurance argument represented in GSN [5].

GSN now offers an accepted way to represent dialectic discourse, as an extension to 'core' GSN. Figure 2:6-1 illustrates how such a dialectic extension might be included within a typical assurance argument using GSN. The initial argument is within the blue box and encompasses the assertions that directly follow from the top goal. The additional dialectic discourse is described below:

- A counter-claim CG1 is added.

- A challenge to the counter-claim CG2, supported by evidence Sn3, defeats the counter-claim CG1.
- Some counter-evidence CSn1 is found.
- A challenge to the counter-evidence CG3, supported by evidence Sn4, defeats the counter-evidence CSn1.

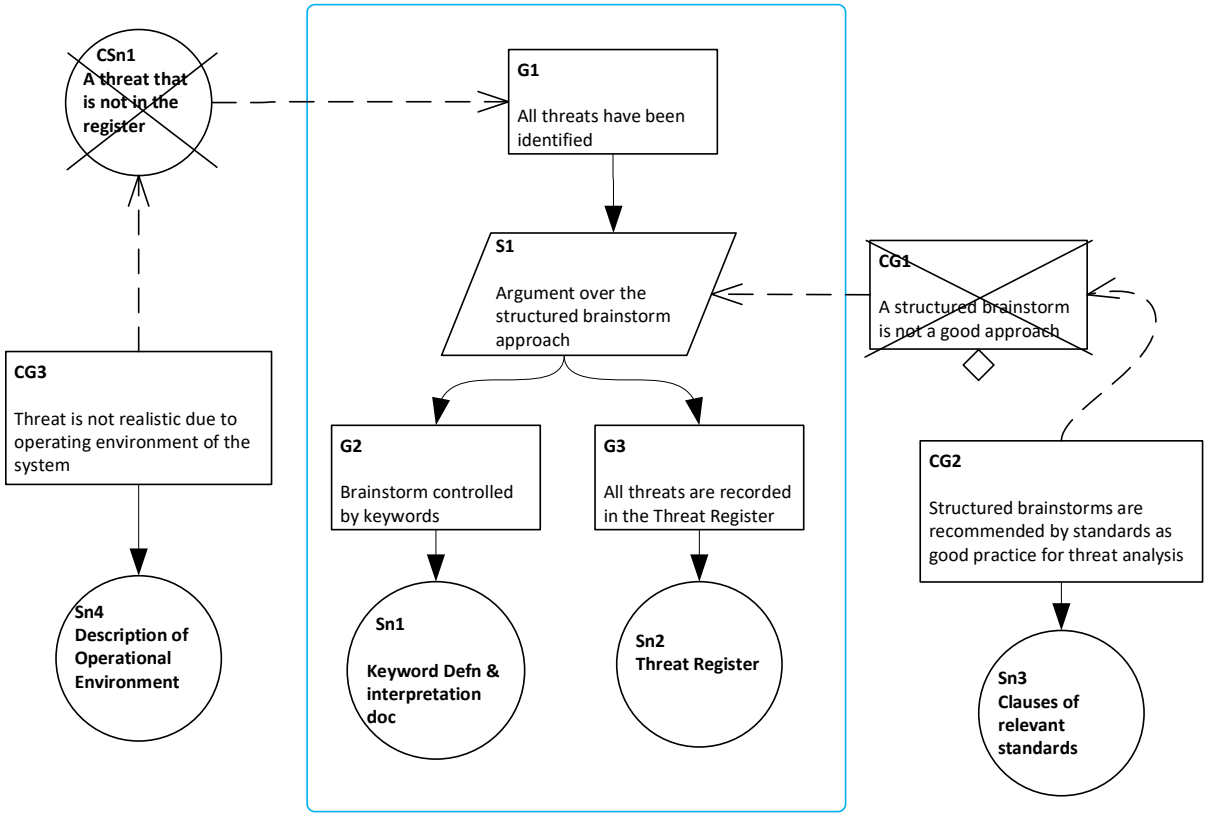


Figure 2:6-1 – GSN Example

2:6.3.5 Contribution to New/Novel Technologies

Novel processes e.g. those related to autonomy, that do not yet have well-established standards and practices will require a more systematic approach to assurance to manage the risks. Along with the increasing pressure for re-use, the current (traditional) approaches to assurance can limit confidence in assurance of such entities. In order to realise the benefits of new technology, a move beyond the capability considerations is needed to assure the process [29].

A recent study as part of a publication on dialectics [30] shows how two real-world accident scenarios chosen at random that involved new/novel technologies, exhibited very similar themes. Both clearly demonstrated the benefit that can be gained from adopting a dialectic approach to robustly challenge the assurance argument, in particular, the potential to reveal at least some of the vulnerabilities of the entities in question, in a timely manner. These vulnerabilities included undue pressure to minimise required assurance, the presence of single point failures, a failure to accommodate changes to the operating environment, and an absence, or poor quality, of operating procedures.

2:6.3.6 Retention of Dialectic Reasoning

One of the challenges of applying dialectic argument is how to present the resulting assurance case when ‘all is said and done’ i.e. the dialectic process has done its job, challenging the argument(s) and forcing updates, including the removal of reliance on untrustworthy evidence.

Much of the value of dialectic reasoning is being able to show the thought process applied, and how the argument has improved by answering the questions posed. This adds to the demonstration of rigour. However, this will not necessarily show in the final argument, as that should be “dialectic-resistant” i.e. resistant to further reasonable challenge, on the basis that all such challenges should have been made during the process of applying dialectic reasoning to the original argument. This is never quite true of course, as such certainty is unrealistic.

In theory it should not matter that this is the case, as the resulting argument should be sound and therefore acceptable. Nonetheless, having a method of avoiding the same questions being applied over again could be considered useful. To that end, a visual representation can be envisaged, to simply and efficiently show where challenges were raised and how they have been resolved. Such presentation is very much down to the preference of the practitioner. Possibilities include that the dialectic elements remain in the final main argument, use of a tool that supports hiding levels of detail, or the use of a side-argument related but separate to the original argument.

The use of a side-argument offers two main benefits:

- Leaving the final, acceptable argument less cluttered and clear to see;
- Retaining the “meta-data” of the dialectic process e.g. for an audit trail.

It is often meta-data that are the most valuable when it comes to any audit activity. The finished artefact, no matter how many iterations it has been through and how carefully its modifications have been tracked, simply cannot present all pertinent information e.g. who reviewed it, when and how those review inputs were resolved. This is especially important where no change to the artefact itself has taken place e.g. a rejected comment, with justification.

Retention of the dialectic reasoning for presentation on demand can be helpful in preserving its value whilst promoting clarity in the delivered argument. This also can be an excellent way of recording the outcome of review or audit and presenting a realistic view of overall confidence in the resulting argument for decision-makers and/or other stakeholders.

2:6.3.7 When to Stop?

The practicality and the efficacy of dialectic argument can be questioned. [30] provides examples that demonstrate how a systematic approach to dialectic argument, in this case directed by guidewords, can be used to establish when enough has been done.

Such application of a full set of guidewords to every part of the argument can be used to provide an end-point to the dialectic activities. The validity of the claim that ‘enough has been done’ relies on the development of a set of dialectic guidewords that are deemed to be complete and this is addressed by the existing work identified in [30].

The nature of a dialectic approach does mean that sometimes the original argument requires changes to address the challenges raised against it. This can then bring a new round of challenge on the revised argument, and hence the potential for a never-ending chain of challenge-update-challenge. It is however reasonable, and recommended, to make the revisions in a manner that pre-empts further challenge, and therefore quickly reach an argument capable of resisting further significant challenge.

2:6.4 Conclusions

This guidance shows what dialectic arguments are, how they should be used and the benefits of doing so.

Assessment of how dialectic argument is used today indicates that it is currently limited in its application, often in both depth and breadth. It has been shown by example that the limited adoption to date can be expanded and the potential benefits of doing so have been demonstrated.

Section 2:6.2.2 cited some common criticisms against assurance cases, which dialectics can address:

Lost confidence in the argument's conclusions;

> Systematic analysis of evidence offered to satisfy argument claims maximises the likelihood of finding such weaknesses at the point of authorship

Logical fallacies;

> In addition to the "stand back" moment to look at the overall argument, examining the detailed way in which specific claims are connected helps to find such fallacies

Insufficient premise in support of the conclusion;

> Specifically addressing the argument context, rather than just evidence (for "counter-examples") helps to ensure the supporting premise is sound

Confirmation bias.

> With the above bias, an argument can be self-fulling in that it is unduly focussed on establishing a stated position to the detriment of establishing the truth. Systematic challenge of the argument and evidence helps to address such bias and so redresses the balance.

Employing dialectic argument to challenge an assurance case has been shown to be highly beneficial. It could be particularly valuable when dealing with emergent technologies where the development and assurance processes are immature and the reduced availability of trusted practices can lead to greater risk. Further benefit could be realised if widespread and well understood sources of risk were to be included in a systematic approach to dialectic argument.

There are still undoubtedly some issues with the adoption of dialectic arguments, not least overcoming the fear of "adding yet more assurance effort" to what is already a potentially under-resourced element of development. This guidance, however, shows systematic challenges to arguments whilst still in the authoring phase has a clear end point and can bring significant benefit. There are also benefits to be gained from a dialectic approach to review. It is hoped that this will lead to a wider take-up of such approaches and will further acceptance as good practice.

2:6.5 Further Information

a) "Computer-assisted Safety Argument Review" [31] presents a dialectical model for assurance (specifically safety) argument review based on a set of "moves", or responses to each part of the safety argument (e.g. a reviewer may present a "counter-argument", a "resolution demand", a "challenge", a "question" and so forth).
   - "A Dialogue-Based Safety Argument Review Tool" [32] formalises this model in a computer tool, evaluated by safety engineers at the University of York. It identifies the elements of a GSN argument (although the elements themselves are notation-agnostic) which may be challenged using a dialectical approach.

b) "Issues Around Dialectic Argumentation" [33] provides a more general discussion on dialectic sufficiency of an argument. This identifies a set of attributes for sufficiency of an argument (coverage, directness, dependence and robustness), with the implication that a dialectical approach may challenge an argument based on one or several of these grounds.
   - This work also identifies two forms of defeat: rebuttal and undercutting, both of which can be applied at any level within an assurance case to challenge a top-level claim or a lower-level sub-claim.

c) “Experience With Assurance Case Preparation” [34] and “Eliminative Induction: A Basis for Arguing Confidence in System Properties” [35] together identify a dialectical approach to authoring (as opposed to reviewing) an assurance (specifically safety) argument. These authors present a formal, structured approach to challenging each step of the safety argument, again using the defeaters rebutting and undercutting, but adding a new one; undermining. Undermining is defined as specifically raising doubts about the validity of evidence, and so can be treated simply as a rebuttal applied specifically to evidence.

d) “Six Honest Serving Men for Today (Dialectic Argumentation for Assurance Cases)”, a published paper seeking to raise the profile of dialectic argumentation within the Safety (and other) Communities, drawing on some of this existing dialectic discussion and suggesting some ways forward with techniques [30].

e) A number of artefacts cover general confidence assessment of assurance (specifically safety) arguments, ranging from theoretical to notation- specific:
   - “The Application of Bayesian Belief Networks to Assurance Case Preparation” [36] and “Towards Measurement of Confidence in Safety Cases” [37] present work on the use of Bayesian Belief Networks.
   - “A New Approach to Creating Clear Safety Arguments” [38] constructs separate confidence arguments.

f) “Confidence: Its role in Dependability Cases for Risk Assessment” [39] seeks to provide confidence by estimating numerical probabilities for evidence.
   - “Eliminative Argumentation for Arguing System Safety – A Practitioner’s Experience” [40] reports on the application of eliminative argumentation to seven different software-intensive systems in the automotive, rail, and industrial control industries. It suggests that the doubt-driven approach to argumentation increases confidence in a case and can be used to support activities such as independent assessments and verification and validation.

PART 3 SUPPORTING INFORMATION

3:1 ACRONYMS, ABBREVIATIONS, DEFINITIONS & GLOSSARY

3:1.1 Acronyms & Abbreviations

Acronym/ Abbreviation	Expansion
ACP	Assurance Claim Points
ACWG	Assurance Case Working Group
ALARP	As Low As Reasonably Practicable
AV	Autonomous Vehicle
CAE	Claims Argument Evidence
GSN	Goal Structuring Notation
HSE	Health & Safety Executive (UK)
MOD	Ministry of Defence (UK)
QALY	Quality-Adjusted Life Year
SACM	Structured Assurance Case Metamodel
SCSC	Safety Critical Systems Club

3:1.2 Definitions & Glossary

3:1.2.1 Assurance Terms

Term	Definition	Source
Argument	A collection of propositions – one of which is the conclusion; the others being the premises for that conclusion	[41]
	Note: This is also taken to include a sequence or chain of arguments where the conclusion of one argument forms a premise to another argument, forming a larger argument.	ACWG
Argumentation	An approach or process of reasoning	ACWG
Artefact	document, record, model or other suitable representation of information Note1: this is used in this guidance in preference to 'document' to avoid any inferred association with capture in a physical form (see also 'artefact'). Exceptions exist for self-reference, quotations and examples. Note2: the guidance recognises that the future of engineering has moved away from past physical document centric ways. The intent is to be neutral in the form that the capture of relevant information will take.	ACWG
Assurance	Grounds for justified confidence that a claim has been or will be achieved	[2]

Term	Definition	Source
Assurance Argument Module	A section of an assurance argument addressing a specific purpose and which has is bounded by an explicit well-defined interface to the remainder of the assurance argument.	ACWG
Assurance Case	reasoned, auditable artefact created that supports the contention that its top-level claim (or set of claims), is satisfied, including systematic argumentation and its underlying evidence and explicit assumptions that support the claim(s) Note: An assurance case 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); • justification of the choice of top-level claim and the method of reasoning.	[2]
Assurance Case Report	A summary document that is designed to track the status of the planned activities, arguments, evidence, risks, opportunities, and issues as defined in the assurance case	ACWG
(Assurance) Entity	The 'thing' that sets the scope of an assurance case. For example: system; product; component; element; facility; service; activity; process; procedure.	ACWG
(Assurance) Property	The property of the assurance entity that is the focus for the assurance case. For example: safety; security; availability. Note: More than one property may be included in the focus for an assurance case. The use of the singular form should not be taken to exclude multiple properties.	ACWG

Term	Definition	Source
Claim	True-false statement about the limitations on the values of an unambiguously defined property (called the claim's property) and limitations on the uncertainty of the property's values falling within these limitations during the claim's duration of applicability under stated conditions Note 1: Uncertainties also may be associated with the duration of applicability and the stated conditions. Note 2: A claim potentially contains the following: • claim's property; • limitations on the value of the property associated with the claim (e.g. on its range); • limitations on the uncertainty of the property value meeting its limitations; • limitations on duration of claim's applicability; • duration-related uncertainty; • limitations on conditions associated with the claim; • condition-related uncertainty. Note 3: The term "limitations" is used to fit the many situations that can exist. Values can be a single value or multiple single values, a range of values, or multiple ranges of values, and can be multi-dimensional. The boundaries of these limitations are sometimes not sharp, e.g. they may involve probability distributions and may be incremental.	[2]
Comprehensive	Including or dealing with all or nearly all elements or aspects of something	[28]
Confidence	The feeling sure of a fact or issue; assurance, certitude; assured expectation	[28]
Conformance	Voluntary adherence to a standard, specification, guide, process or practice. Note 1: 'Conformance' is used in preference to 'compliance', which is forced adherence to a law, regulation, rule or process. Note 2: In this guidance, conformance is taken to include compliance.	ACWG[17]
Conclusion	A judgement or statement arrived at by any reasoning process; an inference, deduction, induction	[28]
Counter-argument	An argument on the opposite side, or against anything	[28]
Counter-evidence	Evidence tending to refute or rebut other evidence	[28]

[17] Derived from https://simplicable.com/new/conformance-vs-compliance

Term	Definition	Source
Dependability	used to describe the availability performance and its influencing factors: reliability performance, maintainability performance and maintenance support performance. Note: Dependability is used only for general descriptions in non-quantitative terms. Note 2: ISO/IEC 25010 notes that “dependability characteristics include availability and its inherent or external influencing factors, such as: reliability, fault tolerance, recoverability, integrity, security, maintainability, durability, and maintenance support.” Several standards address dependability, and many more address the qualities within it. IEC 60050-191 offers related definitions.	[2]
Evidence	Objective artefacts being offered in support of one or more claims	[42]
Inference	Reasoning from something known or assumed to something else which follows from it.	[28]
Premise	A previous statement or proposition from which another is inferred or follows as a conclusion	[28]
Property (see also ‘assurance property’)	An attribute, characteristic, or quality [of a thing]	[28]
	Note 1: A property might include a condition, a characteristic, an attribute, a quality, a trait, a measurement, and a consequence. Note 2: A property might be invariant, or dependent on time, situation, or history	[2]
Record *[verb]*	To relate, narrate, or mention in a written account; to put or set down in writing or some other permanent form; to put on record.	[28]
	Note: this is used in this guidance in preference to ‘document’ to avoid any inferred association with capture in a physical form (see also ‘artefact’). Exceptions exist for self-reference, quotations and examples.	ACWG
Stakeholder	Person or organization that can affect, be affected by, or perceive themselves to be affected by a decision or activity Note: A decision maker can be a stakeholder.	[41]
	Note 1: Stakeholders might benefit from, incur losses from, impose constraints on, or otherwise have a “stake” in the system, and therefore are those that provide the requirements for the system. Note 2: A different but important kind of stakeholder is an attacker, who certainly imposes constraints or has interests involved with the system. Note 3: Depending on conditions and consequences, the various stakeholders require grounds for justified confidence in properties of the system for which they identified requirements.	[2]

Term	Definition	Source
Structured Argument	A particular kind of argument where the relationships between the asserted claims, and from the evidence to the claims are explicitly represented	[7]
[system]	The entity that is the subject of the case. Note 1: Throughout this document the term is used to denote the entity that is the subject of the case. It can be substituted with the relevant term according to context and scope. For example: system; product; component; element; facility; service; activity; process; procedure. See also '(Assurance) Entity' above.	ACWG
Trustworthy	Worthy of trust or confidence; reliable, dependable.	[28]

3:1.2.2 Argument attribute Terms

Term	Definition	Source
Consistent	State where it is possible for all the propositions forming an argument to be true together. (formal logic)	[8]
Deductive	If premises are true, then the conclusion must also be true.	[43]
Inductive	The conclusion follows from the premises not with necessity but only with probability.	[43]
Sound	State where the argument is valid and all its premises are true. (formal logic)	[8]
Valid	State where it is not possible for all of an argument's premises to be true and its conclusion to be false (formal logic)	[8]
	Note 1: the word "valid" does not refer to the truth of the premises or the conclusion, but rather to the form of the inference; Note 2: an inference can be valid even if the premises are false, and can be invalid even if some parts are true; Note 3: a valid form with true premises will always have a true conclusion.	ACWG

3:1.2.3 Risk Terms

These terms are introduced here to give context to the assurance case guidance material. Many of these terms have alternative definitions across different sectors and standards. This guidance material does not seek to impose the definitions below, and in practice the differences in definitions have little effect on the concepts and issues addressed by this guidance document.

Term	Definition	Source
Consequence	Outcome of an event affecting objectives Note 1: An event can lead to a range of consequences. Note 2: A consequence can be certain or uncertain and can have positive or negative effects on objectives. Note 3: Consequences can be expressed qualitatively or quantitatively. Note 4: Initial consequences can escalate through knock-on effects	[41]
Event	Occurrence or change of a particular set of circumstances Note 1: An event can be one or more occurrences, and can have several causes. Note 2: An event can consist of something not happening. Note 3: An event can sometimes be referred to as an “incident” or “accident”. Note 4: An event without consequences (3.6.1.3) can also be referred to as a “near miss”, “incident”, “near hit” or “close call”.	[41]
Likelihood	Chance of something happening Note 1: In risk management terminology, the word “likelihood” is used to refer to the chance of something happening, whether defined, measured or determined objectively or subjectively, qualitatively or quantitatively, and described using general terms or mathematically [such as a probability or a frequency over a given time period]. Note 2: In risk management terminology, “likelihood” is used with the intent that it should have the same broad interpretation as the term “probability” has in many languages other than English.	[41]
Risk	Effect of uncertainty on objectives [of a stakeholder] Note 1: An effect is a deviation from the expected - positive and/or negative. Note 2: Uncertainty is the state, even partial, of deficiency of information related to, understanding or knowledge of, an event, its consequence, or likelihood. Note 3: Objectives can have different aspects (such as financial, health and safety, and environmental goals) and can apply at different levels (such as strategic, organization-wide, project, product and process). Note 4: Risk is often expressed in terms of a combination of the consequences of an event (including changes in circumstances) and the associated likelihood of occurrence. [thus providing the basis for a risk metric]	[41]
Safety	Freedom from risk [of harm] which is not tolerable	[42]

3:1.2.4 Other Terms

Term	Definition	Source
Activity	A collective term incorporating project delivery, system changes and service delivery.	
duty holder	A concept frequently adopted to refer to the person or organisation that is accountable. Note 1: typically the accountability is in relation to a statutory duty but in this guidance it is taken to relate to the property of interest. Note 2: In any given case there can be more than one duty applicable and therefore there can be more than one duty holder. Note 3: The duty holder can be a "virtual person" known as a *persona ficta* in legal terms, such as a corporate body, or may be a "natural person". Note 4: Examples of specific duty holders can be found in [11] (UK military air context) and at https://iadclexicon.org/duty-holder/ (UK Offshore context). Note 5: In some contexts that duty holder may be recognised by another term such as 'Licence Holder' (Nuclear) or 'Accountable Manager' (Aerospace).	
MEDEVAC	Emergency evacuation of the sick or wounded	

3:1.3 Concepts

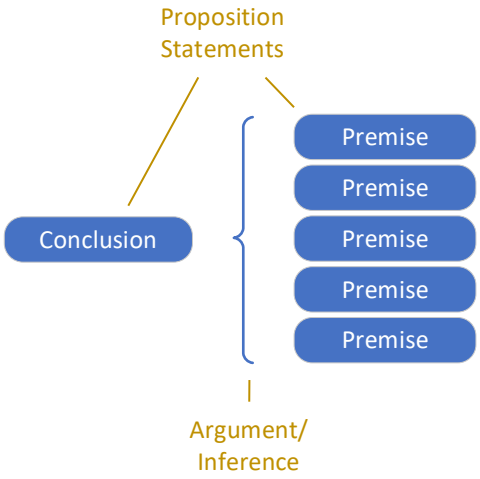
Ancient Greek philosophers defined a number of syllogisms, correct three part inferences that can be used as building blocks for more complex reasoning. This is exemplified in the following famous case often used in introduction to logic:

- Premise: All men are mortal
- Premise: Socrates is a man
- Conclusion: Socrates is mortal

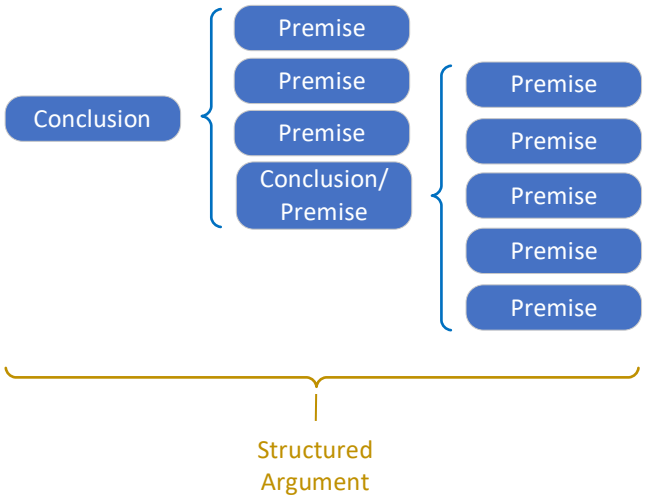
It is noted that each of the premises and the conclusion are propositions, that is they can be evaluated to be true or false. The reader can check that the premises and conclusion are true. Should it be shown that in the context of the argument, 'Socrates' is actually a cat named after the philosopher, one of the premises is false and therefore the argument would be considered 'unsound', even though the logic may be 'valid'.

Arguments where the truth of the conclusion follows definitively from the truth of the premises are said to be deductive, whilst those where the probable truth of the conclusion is inferred from the truth of the premises are said to be inductive.

The concept can be extended to form a conclusion based on any number of premises:



Conclusions from one argument can be cascaded to form the premise of another argument to form a structured argument:



3:2 REFERENCES

The following references are made from this document. It is not intended to reflect and exhaustive set of references on the topic of assurance cases.

[1] The Nimrod Review: An independent review into the broader issues surrounding the loss of the RAF Nimrod MR2 Aircraft XV230 in Afghanistan in 2006. Haddon-Cave, dated 28 October 2009. https://www.gov.uk/government/publications/the-nimrod-review

[2] ISO/IEC 15026-1: 2019; Systems and Software Engineering - Systems and Software Assurance Part 1: Concepts and vocabulary

[3] ISO/IEC/IEEE 24748-1:2018 Systems and software engineering — Life cycle management — Part 1: Guidelines for life cycle management

[4] ISO 9001: 2015; “Quality Management Systems - Requirements”

[5] Goal Structuring Notation Community Standard https://www.scsc.uk/gsn

[6] Claims Argument Evidence Framework https://claimsargumentsevidence.org/

[7] Structured Assurance Case Metamodel https://www.omg.org/spec/SACM

[8] Arguing Safety – A Systematic Approach to Managing Safety Cases; T. P. Kelly University of York, 1998. : https://www.researchgate.net/publication/2565194

[9] White Paper on the Use of Safety Cases in Certification and Regulation; Leveson. http://sunnyday.mit.edu/SafetyCases.pdf

[10] Structured Assurance Cases: Three Common Standards; Scott AT, Krombolz AH. High-Assurance Systems Engineering, 2005. pp. 99–108. https://www.computer.org/csdl/proceedings-article/hase/2005/23770099/12OmNBDgZ0S

[11] MAA02: Military Aviation Master Glossary, Issue 9 https://www.gov.uk/government/collections/maa-regulatory-publications

[12] CAP 670 Issue 3, June 2019. Civil Aviation Authority Air Traffic Services Safety Requirements. www.caa.co.uk/CAP670

[13] Safety Assessment Principles for Nuclear Facilities; 2014 Edition, Revision 1 (January 2020). Office for Nuclear Regulation. https://www.onr.org.uk/saps/index.htm

[14] Judgment Under Uncertainty: Heuristics and Biases; Kahneman, Daniel, Paul Slovic, and Amos Tversky. 1982. New York: Cambridge University Press. ISBN 978-0521284141

[15] Cognitive Biases Potentially Affecting Judgment of Global Risks; Yudkowsky, Eliezer. 2008. New York: Oxford University Press. https://oxford.universitypressscholarship.com/view/10.1093/oso/9780198570509.001.0001/isbn-9780198570509-book-part-9

[16] Thinking, fast and slow; Kahneman, Daniel. 2011. Farrar, Straus and Giroux. ISBN 978-0141033570

[17] UK Health and Safety Guidance on the application of ALARP https://www.hse.gov.uk/managing/theory/alarpglance.htm

[18] The Nuremberg Code [on clinical research] 1947 https://history.nih.gov/display/history/Nuremberg+Code (see also https://history.nih.gov/display/history/Human+Subjects+Timeline)

[19] Airbags Factsheet. The Royal Society for the Prevention of Accidents - Road Safety Factsheet. March 2021 https://www.rospa.com/media/documents/road-safety/airbags-factsheet.pdf

[20] Traffic Safety Facts: Occupant Protection In Passenger Vehicles | 2017 Data. US Department of Transportation - National Highway Traffic Safety Administration (NHTSA) National Center for Statistics and Analysis https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812691

[21] Extract from Hansard - HC Deb 06 April 1977 vol 929 https://api.parliament.uk/historic-hansard/commons/1977/apr/06/speed-limits

[22] Discussion on factors affecting Risk Trade Offs. Science & Technology Committee, Minutes of Evidence. UK House of Commons: HC 428 https://publications.parliament.uk/pa/cm201213/cmselect/cmsctech/428/120201.htm

[23] Preliminary Highway Report HWY18MH010, 2018. US National Transportation Safety Board (NTSB) https://www.ntsb.gov/investigations/AccidentReports/Pages/HWY18MH010-prelim.aspx

[24] SEI Architecture Trade-Off Analysis https://resources.sei.cmu.edu/library/asset-view.cfm?assetid=513908

[25] MSSC 201, Modular Software Safety Case Process Description, IAWG https://github.com/IAWG/MSSC-Process

[26] A New Approach to Creating Clear Safety Arguments; Hawkins, R. Kelly, T. Knight, J. and Graydon P. https://scsc.uk/scsc-109

[27] RTCA DO-178B, Software Considerations in Airborne Systems and Equipment Certification, RTCA, 1992.

[28] Oxford English Dictionary https://www.oed.com

[29] SCSC Workshop on Safety Assurance of Autonomy and AI, University of York, Assuring Autonomy International Programme, 4 February 2019 https://www.york.ac.uk/assuring-autonomy/news/events/scsc-autonomy-workshop/

[30] Six Honest Serving Men for Today (Dialectic Argumentation for Assurance Cases); Oakshott, Y., Chinneck, P. Proceedings of 29th Safety Critical Systems Symposium (SSS'21), February 2021 https://scsc.uk/scsc-161

[31] Computer-assisted Safety Argument Review - A Dialectics Approach; Yuan, T., Kelly, T. and Xu, T. , Argument and Computation, Vol 6, pp 130 - 148, 2015. https://content.iospress.com/articles/argument-and-computation/927921

[32] A Dialogue-Based Safety Argument Review Tool, Yuan, T., Kelly, T., Xu, T., Wang, H. and Zhao, L. 2013. https://www-users.cs.york.ac.uk/ty513/Papers/Yuan et al. AAA2013.pdf

[33] Issues Around Dialectic Argumentation; Kelly, T. Assurance Case Working Group presentation, 2018. https://scsc.uk/file/gc/ACWG5_DialecticArgumentation-440.pdf

[34] Experience With Assurance Case Preparation; Hobbs, C. Blackberry QNX, 2018 https://scsc.uk/file/gc/ExperienceAssuranceCasePreparation-Hobbs-513.pdf.

[35] Technical Report CMU/SEI-2015-TR-005. Eliminative Induction: A Basis for Arguing Confidence in System Properties, Goodenough, J.B., Weinstock, C.B. and Klein, A.Z., Software Engineering Institute, Carnegie Mellon University, 2015. https://resources.sei.cmu.edu/asset_files/TechnicalReport/2015_005_001_434813.pdf

[36] The Application of Bayesian Belief Networks to Assurance Case Preparation; Hobbs, C. and Lloyd, M. Proceedings of the Twentieth Safety-Critical Systems Symposium, 2012. https://scsc.uk/scsc-116

[37] Towards Measurement of Confidence in Safety Cases; Denney, E. Pai, G. and Habli, I. Proceedings of the 2011 International Symposium on Empirical Software Engineering and Measurement, 2011. https://ntrs.nasa.gov/citations/20110016239

[38] A New Approach to Creating Clear Safety Arguments; Hawkins, R., Kelly, T., Knight, J. and Graydon, P. Proceedings of the Nineteenth Safety-Critical Systems Symposium, 2011. https://scsc.uk/scsc-109

[39] Confidence: Its role in Dependability Cases for Risk Assessment; Bloomfield, R., Littlewood, B. and Wright, D. Proceedings of the 37th Annual IEEE International Conference on Dependable Systems and Networks, 2007. https://openaccess.city.ac.uk/id/eprint/1618/

[40] Eliminative Argumentation for Arguing System Safety - A Practitioner's Experience. Published in: 2020 IEEE International Systems Conference (SysCon) https://ieeexplore.ieee.org/abstract/document/9275852

[41] ISO GUIDE 73:2009, Risk Management - Vocabulary

[42] ISO GUIDE 51: 2014, Safety aspects - Guidelines for their inclusion in standards

[43] The Philosopher's Toolkit - A Compendium of Philosophical Concepts and Methods; J. Baggini and P.S. Fosl. Wliey-Blackwell 2010. ISBN 978-1405190183

3:3 CONTRIBUTORS AND ACKNOWLEDGEMENTS

This document and the supporting papers have had the benefit of contributions from a large number of people, who work for a variety of organisations, which collectively span a range of different sectors. Note that contributions have been made on an individual basis and the inclusion of an individual or organisation in the following list does not necessarily mean that individual or organisation agrees with the contents of the document.

3:3.1 Individual Contributors

Antony Edwardson	Mark Carter	Sean White
Andy Scott	Martyn Clarke	Simon Diemert
Chris Hobbs	Matt Osborne	Steve Barrett
Emma Taylor	Paul Chinneck	Stuart Tushingham
Gavin Wilshire	Pete Hutchison	Tim Kelly
Ibrahim Habli	Phil Williams	Yuji Hirao
Jane Fenn	Paul Mayo	Yvonne Oakshott
John Holmes	Philippa Ryan	
Lawrie Henery	Richard Hawkins	

3:3.2 Contributing Organisations

Adelard	Inzpire	RPS Group
BAE Systems	Leonardo	RazorSecure
BlackBerry QNX	NATS	SCSS
Capgemini	NHS Digital	SQEP
Critical Systems Labs	RINA Consulting	The Furious Engineer
Engineer for Safety		University of York

3:3.3 Acknowledgements

This publication was prepared on a consensus basis by the ACWG chaired by Phil Williams. Grateful thanks are extended to all those who participated through preparation of material, its review, or involvement in the lively discussions at ACWG meetings.

ASSURANCE CASE GUIDANCE

This guidance was developed by means of a consensus process involving Assurance Case authors, reviewers, assessors and users from academia and industry. It is thus guidance written for the community, by the community.

Responsibility for publication and maintenance of this guidance rests with the SCSC Assurance Case Working Group (ACWG). See www.scsc.uk/gc for further details.