"EVIDENCE" UNDER A MAGNIFYING GLASS: THOUGHTS ON SAFETY ARGUMENT EPISTEMOLOGY

P.J. Graydon*, C.M. Holloway †

**NASA Langley Research Center, USA, patrick.j.graydon@nasa.gov,*
†NASA Langley Research Center, USA, c.michael.holloway@nasa.gov

Keywords: safety case, assurance argument, epistemology, evidence, safety argument patterns

Abstract

Common definitions of "safety case" emphasize that evidence is the basis of a safety argument, yet few widely referenced works explicitly define "evidence". Their examples suggest that similar things can be regarded as evidence. But the category evidence seems to contain (1) processes for finding things out, (2) information resulting from such processes, and (3) relevant documents. Moreover, any item of evidence could be replaced by further argument. Normative models of informal argumentation do not offer clear guidance on when a safety argument should cite evidence rather than appeal to a more detailed argument. Disciplines such as the law address the problem with a practical, domain-specific epistemology. In this paper, we explore these problems associated with evidence citations in safety arguments, identify goals for a theory of safety argument evidence and a practical safety argument epistemology, propose a model of safety evidence citation that advances the identified goals, and present a related extension to the Goal Structuring Notation (GSN).

1 Introduction

Common definitions of "safety case", like definitions of other forms of assurance cases, emphasize that evidence is the foundation of the safety argument. For example, the *GSN Community Standard* says that a 'reasoned and compelling' assurance argument is 'supported by a body of evidence' [3]. While few commonly referenced works in the field explicitly define "evidence", the examples they give suggest broad agreement that reviews, analyses, and tests are evidence. But there are subtle problems with these definitions and examples. First, the category of things identified as evidence seems as nebulous as the category "software component", including (1) processes for finding things out, (2) information resulting from such processes, and (3) the identity of relevant documents. Second, any item of evidence could be replaced by further argument supported by evidence at a different scope, leading to a potentially infinite regress and a practical question of where to stop. Normative models of informal argumentation do not offer clear guidance on when a safety argument should cite evidence rather than appeal to more detailed argument. Disciplines such as the law address the problem with a practical, domain-specific epistemology. In this paper, we (a) explore the problems of evidence citation in safety arguments, (b) identify goals for a theory of safety argument evidence and a practical epistemology of safety, (c) propose a model of safety evidence citation that advances the identified goals, and (d) present an extension to the Goal Structuring Notation (GSN) to implement our model.

2 "Evidence" in the argumentation literature

In this section, we review how normative texts on safety cases treat evidence, identify problems in that treatment, survey the relevant philosophical literature on the subject, discuss how evidence is treated in other disciplines, and review the treatment of evidence in the recently finalised Structured Assurance Case Metamodel (SACM) [15].

2.1 A process, information, or an artefact?

Normative texts on safety argumentation define evidence both explicitly and implicitly (e.g. through examples). While most experts agree that common safety lifecycle activities such as reviews, tests, and analyses provide evidence, they also define evidence variously as a process, information, or an artefact. These definitions are mutually contradictory and sometimes include things that are not normally thought of as evidence.

The *GSN Community Standard* defines one of two popular graphical notations for recording safety arguments [3]. It defines evidence as 'information or objective artefacts being offered in support of one or more claims'. Assumptions are information and can be offered in support of a claim, but are not evidence. Likewise, a claim might support other claims, but if sub-claims were evidence, the standard's process for top-down construction of goal structures would halt after producing one layer of argument. The standard defines GSN solution elements as 'references to evidence artefacts' [3]. Some of the standard's example solutions, such as 'inspection report', clearly refer to artefacts. Others, such as 'fault tree analysis', might instead refer to a process or information.

Prior to publication of the standard, Kelly's DPhil thesis was the most normative guide to GSN [14]. It does not explicitly define evidence, but it does give many examples of evidence and solutions. Some of these implicitly define evidence as the information produced using a known technique, while others seem to cite processes as evidence. For example, one diagram lists as examples of evidence 'test results, fault trees, and

design information'. Another example cites a 'high quality V&V process' as evidence.

Claims, Arguments and Evidence (CAE) is another popular graphical notation [1, 2]. Adelard's web site defines the CAE evidence element as 'a reference to the evidence being presented in support of the claim or argument, e.g. "the hardware reliability analysis report" [or] "interlock design documentation"' [2]. Guidance from Adelard also notes that 'evidence used at one level of the argument can be: facts, e.g. based on established scientific principles and prior research; assumptions, which are necessary to make the argument, but may not always apply in the "real world"; or sub-claims, derived from a lower-level sub-argument' [1].

2.2 Does evidence end the argument?

In addition to being mutually contradictory and including things such as assumptions, the definitions of evidence cited above share another problem: they do not answer the question of how detailed an argument should be. An argument writer who cannot reliably identify evidence might not know where to begin a ground-up argument or stop a top-down one. One could presumably replace any evidence-supported claim with further argument. For example, rather than supporting a claim that a task has a given worst-case execution time by citing hybrid timing analysis, one could support this claim with an argument over the analysis tool and its inputs [6]. A claim about software behaviour might be supported either directly by test evidence or indirectly by an argument that in turn cites the test results and an analysis of the test plan. The possibility of regress raises the question of where to stop.

Some researchers have noted that details about how reviews, tests, and analyses were carried out is useful in determining how far to trust conclusions drawn from them [11]. But details can obscure the big picture of what it means for a system to be acceptably safe and how it achieves that [9]. As a result, some researchers have proposed presenting some details in a separate confidence argument [11]. But limiting one part of the argument to claims about the system or service in question is not a complete solution; the question of how detailed the confidence argument should be remains.

A developer might ask how much detail an argument should have. One obvious answer is 'as much as your regulator wants to see'. But this answer simply shifts the burden onto the regulator, who might well ask a similar question. The question of what evidence is appears to be crucial, but current answers are insufficient in critical respects.

2.3 Evidence in informal argumentation

These concerns are not unique to assurance arguments. Since assurance arguments use informal logic, one might look to that discipline's literature for answers.

The informal logic initiative began in earnest in the 1970s with the observation that formal deductive logic did not sufficiently equip students to analyse the kinds of arguments they often encountered in the world [13]. The work most often cited as the underpinning of assurance argumentation is Toulmin's 1958 *The Uses of Argument* [17]. Unfortunately, it does not define evidence. While Toulmin uses the word several times, he does so in a way that makes evidence indistinguishable from the more general concept of grounds. Many texts about informal logic have been written since (e.g., [18, 19]). No informal logic text we have read provides a suitable, explicit definition of evidence.

This is not surprising. Informal logic, which overlaps with the pragma-dialectical approach born from linguistics [4], often takes the view that each argument is a dialogue between two parties attempting to determine the truth of a proposition. Because each participant can challenge the other to support a claim, there is no need to distinguish between evidence and further argument. Successful dialogical arguments ultimately rest on propositions that both participants accept as true.

2.4 Evidence in the discipline of law

Safety engineering is not the only discipline that relies on practical arguments. The discipline of law – which inspired Toulmin – is an obvious example, but there are others. For example, the evidence-based medicine initiative seeks to put medical diagnostic and treatment decisions on firmer footing. No discipline we are aware of offers an epistemology that could *mutatis mutandis* solve the problems that concern us here. But, as we will show, other disciplines have found practical substitutes for a sound universal epistemology. This raises hope that the safety discipline might do likewise.

Philosophers have long noted that different disciplines accept different kinds of grounds as sufficient bases for the claims that their practitioners make. For example, Toulmin writes,

> It may turn out … not only that the sorts of grounds to which we point in support of conclusions in different fields are different, but also that the ways in which these grounds bear on the conclusions – the ways in which they are capable of supporting conclusions – may also vary as between fields. There are indications that this may actually be so: e.g. the fact that, though in many cases we speak quite happily of our grounds for putting forward some conclusion as "evidence", in other cases this term would be quite out of place – a man who pointed out the features of a painting which, in his view, made it a masterpiece would scarcely be spoken of as presenting "evidence" that it was a great work of art. [17]

The US and UK legal systems define their own standards of evidence in the form of defined standards of proof (e.g. *preponderance of evidence*, *clear and convincing evidence*, and *beyond reasonable doubt*) and rules of evidence (e.g. the US Federal Rules of Evidence). These are not instances of a perfect, universal theory of knowledge, but instead the result of experts defining and refining a practice with the aim of obtaining the best practical result [20]. In that sense, they embody a *practical epistemology*, a theory of the knowledge relevant to a discipline that can be used to make judgments on relevant questions under prevailing conditions. A practical epistemology is not static. For example, there are proposals to revise the US Federal Rules of Evidence to solve perceived problems [20]. But a practical epistemology, appropriate to

the legal domain, facilitates reasoning with some rigor even in the absence of a sound, practical universal epistemology.

2.5 The Structured Assurance Case Metamodel

In 2013, the Object Management Group (OMG) published the *Structured Assurance Case Metamodel* (SACM) [15]. The bulk of the model, created by the key figures behind GSN and CAE, is intended to facilitate 'collecting, developing, evaluating, communicating, and managing evidence' [15]. By defining what can be documented, the SACM implies aspects of an epistemology. But it is not clear that this epistemology is a suitable practical epistemology for safety.

The SACM purports to 'identif[y] the main factors that determine the evidence collection process, … the main factors that determine the evaluation of evidence, [and] … the elements of evidence' [15]. To this end, it allows arguers to record, among other things, specific items of evidence, the form of evidence (e.g. a document or shell case recovered from a crime scene), relationships between evidence (e.g. this is part of that or belongs to that collection), claims made based on evidence, evaluations of evidence (e.g. whether it supports or challenges a claim, how relevant it is to a claim, the confidence it inspires, its accuracy, the degree of support, the reporting level, and the strength and significance of the evidence), properties of evidence (e.g. its completeness, consistency, reliability, originality, security classification, confidentiality requirements, and version), history and chain of custody (e.g., who did what to it, when, and how), approval and ownership, and the relevant standard of proof.

Much of this seems relevant to understanding how given evidence supports a given claim. But it is not clear that it is practical to record all of this information for each evidence citation in a safety argument. While developers often record some of this information (e.g. configuration history), it has not been shown that recording all of the modelled data will increase the safety of deployed systems at a cost that is not grossly disproportionate. Some model elements seem to have been imported from other disciplines despite having no clear application in safety. For example, the SACM's standards of proof are *unknown*, *other*, *resolved counter evidence*, *beyond reasonable doubt*, *preponderance of evidence*, and *clear and convincing evidence* and it models evidence as either primary or secondary [15]. While the standards of proof are relevant to US and UK jurisprudence and historians routinely classify evidence as primary or secondary, the utility of these concepts in safety argumentation has not been assessed.

3 Goals for a practical epistemology of safety

If we are to build a practice of argumentation around a practical epistemology of safety, we must first define what it means for an epistemology of safety to be practical. In this section, we identify some goals for further consideration.

3.1 Trust, but verify … if you can

An argument is an assertion by its writer: the evidence might provide support or it might not. To assess whether trust in a claim is justified, a reader must either (a) verify that the evidence exists and supports the claim as the arguer says it does or (b) accept the judgment of a capable person who has. A suitable practical epistemology must tell readers how to critically and practically examine the arguments they read.

The capacity of readers to critically analyse a given reasoning step or piece of evidence varies. And it is possible to write arguments that cannot be checked with reasonable effort. For example, consider the claim that given low-level requirements refine specified high-level requirements. This might be supported by (i) appeal to a review of the requirements and traceability matrix, (ii) appeal to such a review with an independent confirmation, or (iii) direct reference to the requirements and traceability documentation. Reading (i) critically requires understanding what makes such a review trustworthy, reading (ii) critically requires knowing what makes an independent review trustworthy, and reading (iii) critically requires undertaking the review. An epistemology of safety defines what it means for readers to read arguments critically. No discipline of argumentation is practical unless the intended audience is capable of the necessary criticism.

3.2 Brevity is the soul of wit

Safety researchers continue to identify information that could be added to safety arguments (e.g. details of execution timing analysis [6]). This trend might be fuelled by a desire to make arguments as close to deductively valid as possible. But detail is added at a cost and that cost is not limited to the cost of recording it: detail presented in an ill-considered way could cause readers to fail to see the wood for the trees [9]. A practical epistemology of safety should facilitate arguing at a level of detail that balances costs and benefits. That is, it should help arguers to write arguments that can be read quickly yet clearly convey both how the system is meant to achieve safety and what is key to its doing so.

Achieving this goal might require abstracting away detail that is relevant yet unlikely to be informative. For example, while a detailed argument over a timing analysis technique might help experts to assess whether its use supports a given claim, it might be more practical to simply cite the analysis as evidence. If the analysis technique can be shown to correctly predict the properties in question, arguing from, not over, the analysis might be briefer, clearer, and just as compelling.

3.3 Sunlight is the best of disinfectants

Regulators or independent assessors might use an assurance argument to identify where the developers' understanding of "adequate" safety, system safety concept, safety plan, safety process implementation, or understanding of techniques and concepts could be improved. Gaps and flaws in the argument might hide or even reflect flaws in the approach to safety assurance for a system. While it is not generally possible to calculate the truth of a safety claim, the process of finding these gaps and flaws might reveal insights that could be used to improve system safety. An acceptable approach to safety argumentation must not hinder the search for such defects by oversimplifying arguments.

3.4 Putting evidence claims on solid foundations

In Toulmin's argument model, *warrants* – often specific to a discipline – are rules that support a given kind of claim based on a given kind of data [17]. It is important to know which warrant supports making a given claim from given evidence for two reasons. The first is that naming the warrant makes it possible to provide backing (justification) just once even if the warrant is used many times in many arguments. The second is that the results that provide backing are to most warrants are defeasible. If new studies show that a given kind of evidence does not support a given claim as previously thought, or does so only in limited circumstances, developers and regulators must identify where the warrant was used and reconsider the continued operation of the affected systems.

3.5 Predictability is the key to managing development risk

While safety arguments discuss *operational risk* (the risk of harm to humans or the environment posed or mitigated by the system in question), developers are also concerned with *development risk* (the risk that the project will fail). The better developers are at predicting the results of independent assessment of their arguments, the lower the risk of rejection and rework. A practical epistemology of safety could help to manage development risk by, among other things, offering clear guidance on how much detail an argument should have.

3.6 Don't paint yourself into a corner

There have been proposals to adopt one or another specific theory of knowledge or confidence in safety arguments. But, as none is known to be perfect, this risks painting the community into a corner. For example, the SACM assumes that confidence can be rated as an integer in the range [0, 100] despite the lack of validation of any means of computing such figures and to the exclusion of alternatives [5].

4 A proposed safety evidence citation model

To help achieve the goals defined in Section 3, we propose viewing the use of evidence in a safety argument as the application of an *evidence scheme* to artefacts to support a claim. Artefacts should be identified precisely, for example by unique identifier, version number, and (if the artefact is a large document) section number. But an artefact alone is not evidence: readers still need to know how it was produced and how to interpret it. Knowing which evidence scheme is being invoked tells them how. Each evidence scheme is a warrant for making a given claim based on given evidence. It:

- Identifies how the artefacts relate to the "evidence", e.g. making clear what test plan documents and test report documents have to do with evidence from testing
- Defines how the evidence is produced and interpreted if applicable, e.g. identifying the kind of review process used to produce source code review evidence
- Tells the reader how to be critical of the evidence, e.g. by identifying a set of critical questions [19] that help to identify challenges to a given instance of the scheme
- Provides a name that links the warrant to established results and discussion surrounding that kind of evidence

Figure 1 illustrates how GSN might be extended to document the application of evidence schemes. We use this example in the subsections below to illustrate how a practical theory of safety knowledge might address the issues raised in Section 3.

The left side of Figure 1 illustrates an appeal to requirements-based low-level testing that achieves Modified Condition/ Decision Coverage (MC/DC) as required by RTCA DO-178C

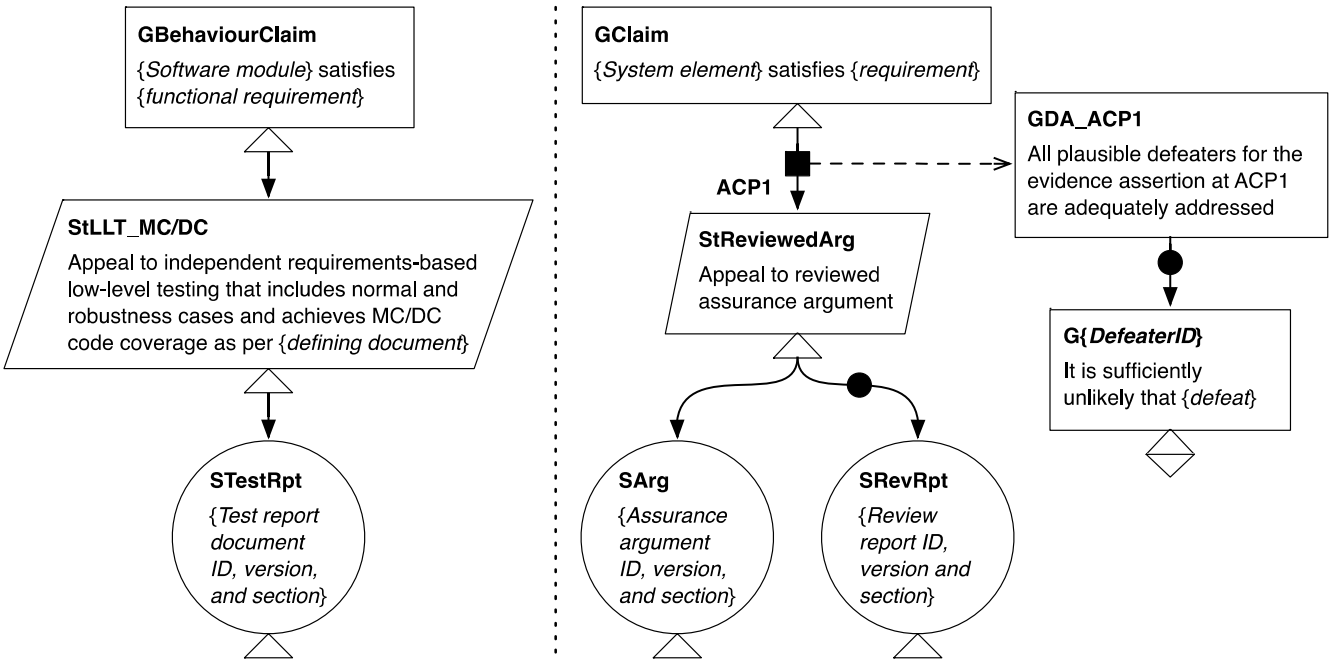

Figure 1: Safety argument patterns for (1) an appeal to a specific form of low-level software testing (left) and (2) an appeal to a separate argument that has been reviewed (right). For a guide to GSN, please refer to the GSN Community Standard [3].

for Level A systems [16]. Goal element **GBehaviourClaim** expresses a claim about a software module, for example that a component for processing raw airspeed sensor data sets a bad data flag if values exceed plausibility thresholds defined by aerospace engineers. Solution element **STestRpt** identifies a relevant artefact, in this case a test report. Strategy element **StLLC_MC/DC** – not normally permitted between a goal and a solution [3] – identifies the evidence scheme. Given the evidence scheme, we do not need to specify whether it is the test report, the results it contains, or some more general information about the testing that is the evidence. Knowing that this reasoning step is the application of the requirements-based MC/DC scheme, readers will know both that the named artefact is a test report and how to evaluate it in conjunction with test plans and other documents that it cites.

The right side of Figure 1 illustrates an appeal to a separate argument or argument module. Goal **GClaim** states a claim about a system element. Solution **SArg** identifies another argument (that presumably concerns the element in question). Solution **SRevRpt** – which might be instantiated many times – identifies a report documenting a review of that argument. Strategy **StReviewedArg** tells us that the writer is appealing to an argument that has been reviewed. While standard GSN does not allow solutions to collaboratively support the same goal [3], the evidence scheme here shows how the other argument and review reports support the claim. If this appeal to evidence occurs in an assured safety argument, **ACP1** and goal **GDA_ACP1** illustrate how the writer could associate a confidence argument with the evidence citation.

4.1 Defining how readers should check evidence

Knowing that **SArg** is being cited as a reviewed argument tells readers how to read it with appropriate scepticism. For example, we might expect the list of critical questions associated with this evidence scheme to include:

- *Does the cited argument make the given claim?*
- *Has the argument been changed since it was reviewed?*
- *Did the review aim to confirm a level of confidence at least as high as what is needed here?*

If the evidence scheme had been an appeal to an *unreviewed* argument, different critical questions would apply. In that case, critical readers might need to review the cited argument.

If the answers to critical questions are not obvious, the writer might provide them in a confidence argument as illustrated in Figure 1. That is, the writer might use the critical questions to enumerate potential argument defeaters and explain why selected defeaters do not, in fact, defeat the argument.

4.2 Promoting compact, focused argument

Testing, like many things, might be done well or poorly. We could expand the argument to discuss how the details of a specific set of tests show that those tests are adequate to support the given claim. But that is a lot of detail, especially where similar testing supports *many* similar claims. Naming the evidence scheme rather than adding detail to the argument keeps the readers' attention focused on the safety-relevant behaviour claim while linking to general knowledge about the sort of evidence in question, including how strong it should be expected to be. The limited set of readers who will be especially critical, such as independent safety assessors, can probe more deeply into the evidence citation as needed.

Previously proposed approaches for keeping safety arguments brief and focused have called for details about evidence to be recorded in separate, linked arguments [8, 11]. But, in the absence of a good absolute metric for confidence in safety claims [5], readers must read the linked argument to know how far to trust a claim supported by evidence. An evidence scheme, in contrast, stands broadly for the evidence-related process and represents a named amount of confidence (even if that amount cannot be represented on any absolute scale). An arguer is free to omit details that can be inferred and so focus the reader's attention only those of special relevance.

4.3 Facilitating effective argument criticism

Evidence schemes facilitate effective argument criticism by associating evidence citations with advice on how to critically examine the given kind of evidence. This advice might take the form of a list of critical questions, but could also take the form of a standard procedure for auditing the kind of evidence in question. Such advice can never be known to be complete. But as the community of engineers, assessors, regulators, and scientists discover new ways in which a given kind of safety evidence might fail to support a given kind of claim, the evidence scheme's name provides a way of linking those lessons learned to the related parts of past and future safety arguments.

4.4 Towards sound warrants for evidence-based claims

Naming evidence schemes provides a means of associating reasoning steps with general – domain-wide if not universal – backing for their warrants. Because each evidence scheme is specific to a claim, a type of evidence, and the processes of producing and interpreting that evidence, it effectively defines research questions relating to the strength of the warrant. For example, the testing-related warrant illustrated in Figure 1 frames research questions about the efficacy of requirements-based testing that achieves MC/DC:

- *If a software module has been successfully tested in this way, how likely is it to meet its functional requirements?*
- *Under which circumstances is this more or less likely?*

To the degree that science answers such questions, it does so with a variety of evidence: pilot studies (perhaps in the form of small case studies), studies of feasibility, experiments (with more or fewer participants who represent the population of interest to varying degrees), historical studies (of the causes of specific accidents or of the performance of a broad class of systems), and so on. No single study answers the question. This makes it difficult for the writer of a safety argument to cite a justification for using a given type of evidence. But as scientists continue to learn more about each type of evidence, regulators can take what is known into account in deciding the circumstances in which to accept each evidence scheme.

Just as the courts have built a practical epistemology in the form of precedent, standards of proof, and rules of evidence, the safety community can build a practical theory of safety knowledge around decisions about the circumstances under which the use of each evidence scheme is permissible or not.

4.5 Making certification more predictable

Rework (and the associated delay) is expensive. As a result, making certification more predictable is key to reducing development risk. While writers of safety arguments in some domains have the benefit of a standard that specifies the evidence their arguments should cite [12], others are faced with the question of where to stop. A practical epistemology of safety – complete with evidence schemes and rules for when each may be used – could provide the needed guidance. A writer using a top-down method proceeds until claims can be supported by applying a known evidence scheme that the relevant regulator approves of, or at least one that has gained community recognition and not yet been discredited.

4.6 Keeping our options open

The decision to document and use named evidence schemes would not tie the assurance argumentation community to any particular theory of evidence or confidence. For example, unlike the SACM, the idea of an evidence scheme does not limit users to specific standards of proof or preclude assessing confidence using eliminative induction [7]. Our proposal seems to require only the ability to name a rule of inference.

5 Conclusion

There is broad agreement in the safety community that safety arguments are based on evidence and that review, test, and analysis provide evidence. But there is little agreement about what evidence is. Information, artefacts, and the processes of generating and interpreting artefacts all play a role, but a precise definition remains elusive. What should and should not be cited as safety evidence is also unclear.

Neither the philosophy literature nor other disciplines that use argument seem to offer a universal theory of knowledge that is applicable to safety arguments. But other disciplines model another way to make reasoned decisions: the explicit creation and upkeep of a practical epistemology. We hypothesize that recognition of a set of rules for what counts as sufficient evidence for a given kind of claim under given circumstances would provide developers, assessors, and regulators with a practical means to make justified decisions about how much detail an argument should have and whether an argument is sufficiently compelling.

Because no practical epistemology is perfect, some decisions made by applying it will be incorrect. As our community develops new techniques and learns more about existing ones, we will need to devise, refine, replace, and repudiate evidence schemes. Nevertheless, using warrants based on our current knowledge might be the best way to keep safety arguments on as sound an epistemic footing as practical.

References

[1] Adelard. *ASCAD: Adelard safety case development manual*. Electronic document, London, UK, 1988.

[2] Adelard. "Adelard - ASCE". Web page: http://www.adelard.com/asce/choosing-asce/cae.html. Last accessed 2 June 2015.

[3] K. Atwood et al. *GSN community standard.* Version 1, Origin Consulting Ltd., UK, Nov. 2011.

[4] F. H. van Eemeren, R. Grootendorst. *A systematic theory of argumentation: The pragma-dialectical approach*. Cambridge University Press, Cambridge, UK, 2004.

[5] P. J. Graydon. "Uncertainty and confidence in safety logic," in *Proc. Int'l System Safety Conf. (ISSC)*, 2013.

[6] P. J. Graydon, I. Bate. "Realistic safety cases for the timing of systems", *The Computer Journal*, 57(5), pp. 759–774, 2014.

[7] J. B. Goodenough, C. B. Weinstock, A. Z. Klein. "Eliminative induction: A basis for arguing system confidence", In *Proc. Int'l Conf. on Software Engineering (ICSE)*, 2013.

[8] I. Habli, T. Kelly. "Achieving integrated process and product safety arguments", In *Proc. 15th Safety-Critical Systems Symp. (SSS)*, 2007.

[9] C. Haddon-Cave. *The Nimrod review: An independent review into the broader issues surrounding the loss of the RAF Nimrod MR2 aircraft XV230 in Afghanistan in 2006*. The Stationary Office, London, UK, 2009.

[10] R. Hawkins, T. Kelly. *A software safety argument pattern catalogue*. Technical report YCS-2013-482, University of York, York, UK, 2013.

[11] R. Hawkins, T. Kelly, J. Knight, P. Graydon. "A new approach to creating clear safety arguments", In *Proc. 19th Safety-Critical Systems Symp. (SSS)*, 2011.

[12] ISO 26262. *Road vehicles — Functional safety*. Int'l Organization for Standardization, 2011–2012.

[13] R. H. Johnson. "Some reflections on the informal logic initiative", *Studies in Logic, Grammar and Rhetoric*, 16(29), pp. 17–46, 2009.

[14] T. P. Kelly. *Arguing safety — A systematic approach to managing safety cases*. DPhil thesis, University of York, York, UK, 1998.

[15] *Structured assurance case metamodel (SACM)*, Ver. 1.0. Object Management Group (OMG), 2013.

[16] RTCA DO-178C. *Software Considerations in Airborne Systems and Equipment Certification*. RTCA, Inc., 2011.

[17] S. E. Toulmin. *The uses of argument*. Cambridge University Press, New York, updated ed., 2003.

[18] D. N. Walton. *Methods of argumentation*. Cambridge University Press, New York, 2013.

[19] D. Walton, C. Reed, F. Macagno. *Argumentation schemes*. Cambridge University Press, Cambridge, 2008.

[20] D. Walton, N. Zhang. "The epistemology of scientific evidence", *Artificial Intelligence and Law*, 21(2), pp. 173–219, 2013.