Quantifying Uncertainty in Safety Cases Using Evidential Reasoning

Sunil Nair[1], Neil Walkinshaw[2], and Tim Kelly[3]

[1]Simula Research Laboratory, Norway
[2]Department of Computer Science, University of Leicester, United Kingdom
[3]Department of Computer Science, University of York, United Kingdom
sunil@simula.no, n.walkinshaw@mcs.le.ac.uk, tim.kelly@york.ac.uk

Abstract. Dealing with uncertainty is an important and difficult aspect of analyses and assessment of complex systems. A real-time large-scale complex critical system involves many uncertainties, and assessing probabilities to represent these uncertainties is itself a complex task. Currently, the certainty with which safety requirements are satisfied and the consideration of the other confidence factors often remains implicit in the assessment process. Many publications in the past have detailed the structure and content of safety cases and Goal Structured Notation (GSN). This paper does not intend to repeat them. Instead, this paper outlines a novel solution to accommodate uncertainty in the safety cases development and assessment using the *Evidential-Reasoning approach* - a mathematical technique for reasoning about uncertainty and evidence. The proposed solution is a bottom-up approach that first performs low-level evidence assessments that makes any uncertainty explicit, and then automatically propagates this confidence up to the higher-level claims. The solution would enable safety assessors and managers to accurately summarise their judgement and make doubt or ignorance explicit.

Keywords: safety, safety assessment, safety case, confidence argument, evidence, evidential reasoning, human factors, expert judgement, uncertainty, confidence

1 Introduction

Goal-based system safety standards such as DS 00-56 (MoD 2004a) often require the construction and provision of a safety case - *a structured argument, supported by a body of evidence, that provides a compelling, comprehensible and valid case that a system is safe for a given application in a given environment* [1]. The assessor needs to establish confidence that the safety case adequately addresses the identification and mitigation of hazards. Unfortunately, both evidence and argument will typically be imperfect and uncertainties in the assessment of safety cases are unavoidable.

A major challenge in developing a good safety case is to determine what type of evidence and how much of this evidence is *sufficient* to satisfy the safety case claims. Expert judgement plays a vital role in this process. However, the developer or the assessor can never be 100% certain that all hazards were mitigated. Furthermore, uncertainties might exist from secondary issues, such as who created the safety case,

who was responsible for generating the evidence, what types of tools and techniques were used, etc. These confidence factors often tend to be implicit considerations in the development and assessment of safety cases.

This paper proposes a novel approach to explore these factors and provide a mathematically sound framework for assessing safety cases using Evidential Reasoning (ER) [2]. The uncertainty of the expert´s judgement is captured in this approach through a series of questions (specific for different evidence types), gauging their confidence in the supporting evidence. ER provides a mechanism by which this low-level confidence information can be propagated up the hierarchy of a structured safety case represented in GSN. The ER algorithm [2] allows us to calculate an aggregate *belief function* for the top-level claim, which explicitly captures any uncertainty in the expert´s judgement from the lower-level confidence ratings. Eliciting the expert´s confidence factors for different evidence types and providing a scale of uncertainty, will allow both developers and regulators to more accurately summarise their opinion and make any doubt or ignorance explicit. This assessment framework will help safety case assessment to be more systematic and consistent, thereby providing increased assurance on the safety of the system.

The rest the paper is organized as follows. Section 2 outlines the background of the paper. Section 3 presents the research agenda and proposed solution. Section 4 presents our conclusions.

2 Background

This section introduces the background on expert judgement in safety and ER. We also review related work.

2.1 Confidence and Uncertainty in Safety Assessment

Recent studies have shown that determining the confidence in the safety of a system as a whole and, as a part of that process, confidence in individual pieces of evidence is challenging for both industry [3] and academia [4]. The strong reliance of judgment-based processes has led to the current situation where expert judgment may be considered as a *de facto* method for assessing safety of a system in practice [3]. Despite the pervasive and predominant use of expert judgment in safety assessment, few systematic investigations on handling uncertainty have been performed to date.

Improving safety case development and argumentation has been a major research interest in the past. The notion of confidence arguments and assurance deficits were introduced to support the safety case development [5]. Studies have also dealt with confidence factors and criteria used in safety assessments [6, 7]. Past studies [8-10] have detailed the notion of uncertainty in safety cases and provided ways to handle them e.g., using Bayesian Belief Networks (BBN) [11]. Although plausible, BBN rely heavily on their probability tables, which in turn rely on the availability of prior probability information. This reliance upon the prior probability information, which is often complicated to obtain, makes it difficult to provide a thorough assessment on confidence where the assessor is ignorant or doubtful.

2.2 Evidential Reasoning (ER)

The general challenge of reasoning about multifaceted decision problems, where the underlying data is subject to varying degrees of (un-)certainty is well-established. In the late 60´s, Dempster and Schäfer proposed that the subjective beliefs of individuals could be expressed as `belief functions' in their `Theory of Evidence' (DS-theory) [12, 13]. In a belief function, the possible range of beliefs is represented as a Likert-scale (e.g., 0 is very bad and 5 is excellent), and the subjective belief is represented as a distribution over this scale (where total ignorance is represented as an empty function). They then showed how such belief functions could be combined to yield aggregate beliefs for multi-faceted decision problems.

In reality (e.g., the assessment of safety cases), decision problems tend be structured; certain factors may feed-in to each other, and can form more complex, hierarchical belief structures. ER [2] is an extension of DS-theory that enables the aggregation of belief functions, where the factors are arranged in a hierarchical structure. The root-node represents the final decision one wishes to make. Branch nodes represent contributory factors. Branches can be given different weights, indicating the extent to which they contribute to the overall decision. Leaf-nodes represent points at which one can present ones own belief functions. ER then provides the mathematically sound basis by which to combine the belief-functions provided in the leaf-nodes, and to propagate them up to the root.

3 Research Agenda

Our overall goal is to develop a tool-supported framework to improve and support expert judgment in safety assessments. Following on from preliminary work using ER to assess software quality [14], we intend to apply ER to provide an automated, mathematically sound basis for the assessment of the expert's confidence in safety-claims, as set out with confidence arguments.

The proposed high-level procedure is shown in Figure 1. A typical GSN confidence argument will allow structuring of claims and the supporting arguments that increase confidence. The satisfaction of the low-level claims relies on the solution (evidence) supporting them. Through a series of generic and specific questions about the solution, the expert will set out their assessment (ranging from a scale of 0 – 5) and their confidence (a quantified value of confidence level e.g., in percentage) in the satisfaction of the claim. ER will then propagate these beliefs through the GSN structure to yield an overall assessment of the system. Crucially, any ignorance or uncertainty about a claim will be made explicit in the overall assessment as well. Some sample generic questions are shown in the Figure 1.

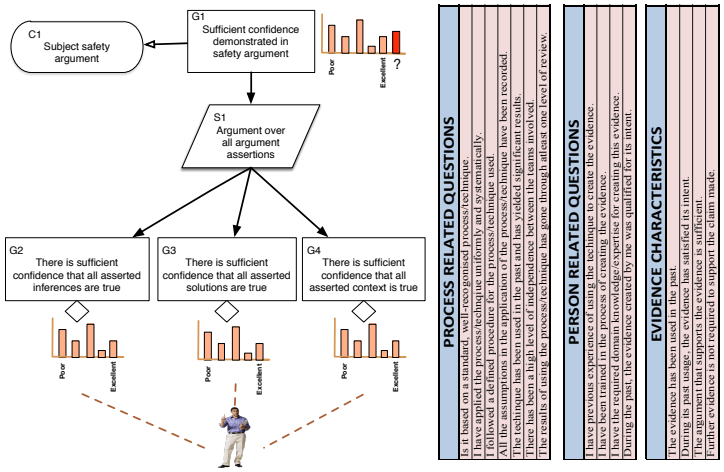

Figure 1. High-level application of ER on GSN confidence argument and sample questions

To achieve the above-mentioned goal we formulate will require us to address the following key research questions (RQs):

RQ1. *What information makes experts gain or lose confidence in the claim, its arguments, and the supporting evidence?*

This RQ aims to understand the expert´s decision-making process. It attempts to identify the various factors and criteria for individual evidence types that influence the confidence of the expert. The key challenge here would be to identify through systematic examination the specific questions to establish the underlying belief functions in ER. Different evidence types are likely to have specific factors that influence the expert´s confidence and these needs to be identified. An initial attempt to answer this RQ was through interviews with experts [10].

RQ2. *How can the confidence in goal structured safety cases be quantified along with uncertainties with the help of ER?*

This RQ aims at adaptation of ER approach to a goal structured safety case. We need to identify ways in which the confidence can be quantified in the argument patterns proposed [5]. We also need to identify ways in which the assurance deficit is captured and communicated to the assessor. As a potential challenge, we need to account for the fact that the confidence arguments are not necessarily tree-structured. We need to identify an approach that enables feeding the answers to the questions (see RQ1) into the ER framework and efficiently propagate these lower-level belief values to the top-level claims. Implementation of the approach with a scalable tool support is also a major step in this RQ.

4 Conclusion

This paper has introduced our position in relation to a potential assessment framework that enables quantification of uncertainties and confidence in safety case with the help of the Evidential Reasoning. The framework enables assessors and

developers to explicitly quantify any ignorance or doubt they have in the assessment of the lower level solutions. The ER algorithm will propagate these confidence values as *belief functions* to the top-level claims while maintaining the GSN structure. Our preliminary investigations on safety case assessment have shown the importance of identifying and building confidence arguments to support the core safety argument and effectively quantify the confidence and the assurance deficit. This will greatly improve the clarity and consequently the comprehension of the arguments and help reduce the overall size of the core argumentation.

We plan to take some initial steps towards answering the research questions in the near future by systematically identifying factors that influence expert´s confidence and how to elicit them. Initial steps have already been taken towards this objective [10]. We also plan to implement the framework as a scalable tool that enables safety case development using GSN and assessment through the adaption of ER on confidence arguments. The tool support would be validated with experts to identify its usefulness in practice. This task would require collaboration among system suppliers and safety assessors in order to investigate the potential of the proposed approach.

Acknowledgments. The research leading to these results has received funding from the FP7 programme under grant agreement n° 289011 (OPENCOSS) and from the Research Council of Norway under the project Certus SFI.

References

1. Interim Defence Standard 00-56 Part 1 - Issue 5, in, UK MOD (2014)
2. Yang J.-B., Xu D.-L.: On the evidential reasoning algorithm for multiple attribute decision analysis under uncertainty, IEEE Transactions on Systems, Man, and Cybernetics, Part A, vol. 32, no. 3 (2002)
3. Nair. S , et al.: The State of the Practice on Evidence Management for Compliance with Safety Standards, Simula Research Laboratory, Techincal Report (2013)
4. Nair. S, et al.: An Extended Systematic Literature Review on Provision of Evidence for Safety Certification, in Information and Software Technology, 56(7), 689-717 (2014)
5. Hawkins. R, et.al.: A new approach to creating clear safety arguments. In Advances in Systems Safety (pp. 3-23) (2011)
6. Hamilton. V.: Criteria for Software Evidence, Goal-based standards require evidence-based approaches, Safety Systems, 16:1 (2006)
7. Nair. S, et.al.: Understanding the practice of Safety Evidence Assessment: A Qualitative Semi-Structured Interview Study, Technical report, Simula Research Laboratory (2014)
8. Denney. E, Pai. G.: A lightweight methodology for safety case assembly. In Computer Safety, Reliability, and Security (pp. 1-12). Springer Berlin Heidelberg (2012).
9. Weaver, R., et.al.: Gaining confidence in goal-based safety cases. In Developments in Risk-based Approaches to Safety (pp. 277-290) (2006)
10. Ayoub, A., et.al.: A systematic approach to justifying sufficient confidence in software safety arguments. In *Computer Safety, Reliability, and Security* (pp. 305-316) (2012)
11. Denney. E, et.al.: Towards measurement of confidence in safety cases. In ESEM (2011)
12. Dempster A. P.: A generalization of Bayesian inference, Journal of the Royal Statistical Society, Series B, vol. 30, pp. 205–247 (1968)
13. Shafer. G.: A Mathematical Theory of Evidence. Princeton University Press, (1976)
14. Walkinshaw. N.: Using evidential reasoning to make qualified predictions of software quality. In PROMISE (2013)