Safety Demonstration – A Strategy for Assessors

André A. Hauge[a], Vikash Katta[a], Peter Karpati[a] and Bjørn Axel Gran[a,b]
[a] Department of Risk, Safety and Security, Institute for Energy Technology, Halden, Norway
[b] NTNU, Trondheim, Norway

Abstract: This paper presents a strategy supporting the planning and documenting of an assessor investigation of whether a Digital Instrument and Control (DI&C) system or system change is sufficiently safe for its intended purpose on the basis of its associated documentation. With assessor, it is meant a project independent assessor of a vendor, a utility, an independent third party, or a regulatory body. The strategy is supported by: (1) a process supporting the planning and documenting of an assessment; and (2) a language for expressing claims, arguments and evidence relationships as well as assessment results. An example case is used in the paper to explain how an assessor may choose and alter the assessment strategy during an investigation and organise evidences as support in order to arrive at a conclusion. The intention is to facilitate the assessor in describing how evidences is organised and evaluated systematically although the assessment process and methods applied may need to be adapted as new knowledge about the target is acquired during the investigation.

Keywords: review, assessment, argumentation, decision process, safety demonstration.

1. INTRODUCTION

The OECD Halden Reactor Project (HRP) has for the last seven years been performing research on safety demonstration of DI&C. The project has investigated the practices, challenges and needs related to safety demonstration by conducting workshops with industrial experts [1][2], in-depth interviews with nuclear regulatory and utility organisations [3], as well as case studies [4][5]. Assessing the DI&C of a new Nuclear Power Plant (NPP), or whether a utility proposed change within a safety critical or safety related system is sufficiently safe is challenged by the large quantity of associated documentation to be reviewed. Furthermore, often it is not possible for the assessor to perform a complete assessment of all aspects of the system due to resource constraints. Therefore, the assessor needs a strategy for limiting the investigation within the resource constraints, while at the same time assuring sufficient coverage in order to obtain the needed confidence. In general, an assessment strategy is shaped by established practices, the experience of the assessor, the type of system and functionality that is affected by the proposed change, the results of relevant and already performed assessments, and the applicable regulatory requirements that the assessor is required to consider.

In this paper, we present a strategy supporting an assessor in planning and documenting an investigation of whether a DI&C system or system change is sufficiently safe for its intended purpose on the basis of its associated documentation and artefacts. In short, the strategy supports the assessor in planning and documenting his or her assessment. With assessor, it is meant a project independent assessor of a vendor or a utility, an independent third party, or a regulatory body. The strategy is supported by: (1) a process supporting the planning and documenting of an assessment; and (2) a language for expressing claims, arguments and evidence relationships as well as assessment results.

2. BACKGROUND

2.1. Relation to state-of-the-practice

On the basis of in-depth interviews with nuclear regulators and technical support organisations [3], some of the main challenges for the DI&C safety assessor are how to:

- Decide how to allocate resources so that issues with high impact on safety is prioritised before issues with low impact and decide what not to assess.
- Decide what kind of methods and tools that are best suited for generating the necessary evidence or evaluate presented evidence.
- Decide when and how to adapt the assessment strategy as new knowledge of the target is acquired.
- Decide what is an acceptable evidence in support of different claims and what is not.
- Decide how to weight different kinds of evidences.
- Decide when to stop the assessment, in the sense determine that the necessary confidence that enables to conclude is acquired.

Although efforts have been made on reaching a common understanding between nuclear regulators [6], there is no single common approach within the nuclear community for how a licensee shall demonstrate safety or how a regulator shall perform its assessment [3], there are several. With safety demonstration, it is here meant [6] "*The set of arguments and evidence elements which support a selected set of claims on the safety of the operation of a system important to safety used in a given plant environment"*. In the following, the main approaches for safety demonstration extracted from interviews with nuclear regulators are outlined. The outlined approaches are not mutually exclusive, the intent is to indicate the main differences which are [3]:

- Compliance approach – A licensee prepares its documentation that contains a description of the fulfilment of detailed requirements issued in regulations and guidance. The regulator assesses the fulfilment of the requirements. The success depends on the ability to demonstrate compliance.
- Goal setting approach – A licensee prepares its documentation and describes the objectives to be achieved and how these are met. The regulator assesses the documentation. The success depends on the ability of the licensee to clearly define the objectives to be met and demonstrate that these are met.
- Safety Case approach – A licensee prepares a safety case as part of the safety documentation. The safety case explicitly describes the claims, arguments and evidences that assures safety. The regulator assesses the safety case. The success depends on the ability to argue the case that safety is assured.
- Project documentation review approach – A licensee prepares the documentation related to a project for review. The regulator assesses the different parts of the project documentation systematically. The success depends on the ability to clearly document the project in such a manner that convinces the assessor, independent on what the assessor chooses to focus on in the investigation, that the system is safe. The licensee is not expected to provide a prepared safety demonstration, the documentation of the project as a whole is within the scope to be considered by the regulator.

2. 2. Relation to state-of-the-art

An important part of a safety demonstration assessment strategy is how an assessor documents his/her take on the presented demonstration. It can be done on a specific basis (e.g. by using a self-defined schema or one offered by the assessor's organisation), or a general basis by utilising notations for depicting arguments. Two main argument notations are the Goal Structuring Notation (GSN) [7] and the Claims, Arguments and Evidence (CAE) notation [8]. According to [7], GSN "is a graphical argumentation notation that can be used to document explicitly the individual elements of any argument (claims, evidence and contextual information) and, perhaps more significantly, the relationships that exist between these elements (i.e. how claims are supported by other claims, and ultimately by evidence, and the context that is defined for the argument)." GSN is a mature and active topic of research. The CAE notation is less well described than GSN, but often used as support for documenting safety cases [9][10][11]. Both notations are supported by tools. One of the most referred

argumentation models is Toulmin's model of argumentation [12], which inspired the GSN and the CAE notations as well as other adaptations as support for developing trust cases [13].

Another important part of the assessment strategy is to address how the assessors builds confidence in in safety being sufficiently demonstrated during the assessment. Different means to assist the process of confidence building is addressed in the literature. Greenwell et al. offers a list of common argument fallacies [14] that can be used as assessment support. Hawkins et al. [15] present an approach called assured safety arguments and suggests that the confidence argument should be defined separate from the argument concerning safety behaviour (hazard management argument). However, the elements of the different kinds of arguments are interlinked in the argument structure such that the confidence argument justifies the sufficiency of confidence in the hazard management argument. Nair et al. [16] builds upon the assured safety arguments approach and presents how the individual judgements of an assessor concerning confidence factors of the evidence (e.g. trustworthiness and appropriateness) can be documented using Likert-scales. In [17], the Trust-IT framework is presented that enables the assessors to express their attitudes by the use of a decision scale and a confidence scale. The decision scale concerns attitude towards acceptance or rejection of an assessed element with the four values: acceptable; tolerable; opposable; and rejectable. The confidence scale concerns confidence in the decision being made with the following six values: for sure; with very high confidence; with high confidence; with low confidence; with very low confidence; and lack of confidence.

2.3. Relation to PSA/PRA

An assessor investigating whether a proposed DI&C system or system change is sufficiently safe for its intended purpose concludes on the documented evidences that risk is or is not sufficiently reduced. Probabilistic Safety Assessment (PSA)/Probabilistic Risk Assessment (PRA) is within the nuclear industry an established technique to quantify risk. Although the PSA/PRA of the DI&C may be a valuable "evidence" supporting the assessor in concluding, the grounds for which the PSA/PRA is based and how these are combined still needs to be evaluated, our strategy supports such an evaluation.

A DI&C system is complex, so is the development process. Although it is possible to quantify hardware failure rates, human performance in operating scenarios and much more, some aspects of the development and operation of a DI&C is not that easily quantified. The main functionality of a DI&C is realised in the software. It is debatable whether quantifying the probability of software failure makes sense, the opponent view being that software is a mathematical and logical construct and thus behaves and fails systematic. Then, in order to reduce risk in the context of the systematic nature of software, two main ways is to assure and provide evidence for an acceptable development process being applied (e.g. following applicable standards and guidelines) and that an acceptable product is achieved in the form of a deterministically safe software. The assessor is then faced with a number of soft evidence (e.g. compliance to an acceptable process) and hard evidence (e.g. proofs of deterministic behaviour) that needs to be evaluated. Our strategy is tailored to support an assessor in organising and evaluating these evidences. The relation of our strategy to PSA/PRA is not a direct one, it is indirect in that it can be used to qualitatively investigate the grounds within the safety demonstration, the same grounds that also may be part of PSA/PRA. In addition, our strategy supports the assessor in quantitatively expressing his or her confidence in the evidences. This numerical value, representing the assessor confidence in the evidences, is subjective and is only intended as support for the assessor in concluding.

3. THE STRATEGY EXPLAINED

On the basis of the main challenges and common approaches for safety demonstration and assessment presented in Chapter 2, we see the need for a flexible assessment strategy. In addition, we claim that there is a need for transparency in how an assessor plan, execute, document and conclude on the basis of the decisions taken and the findings from the assessment. With transparency, it is here meant the

explicit description of the logical structure of claims, arguments and evidences in an argument structure as well as the explicit description of the assessor´s decisions generated during the assessment.

Instead of crafting a strategy for assessors that limits what approach to apply, we rather acknowledge that there are several and a mix approaches that an assessor may apply when organising evidences and assessing them in order to support a conclusion. Our strategy for assessors is expected to be flexible, facilitating the assessor in crafting and adapting its strategy as new knowledge is acquired during the assessment. However, although different assessors may apply different approaches, it is important to have justifiable grounds for the conclusion being right. In order to capture on what grounds an assessor justify its conclusion, our strategy is supported by a language for expressing assessor decisions explicitly.

In the following sub-sections, we describe the process and language that jointly represents the building blocks for an assessor to create and execute a strategy for safety demonstration assessment.

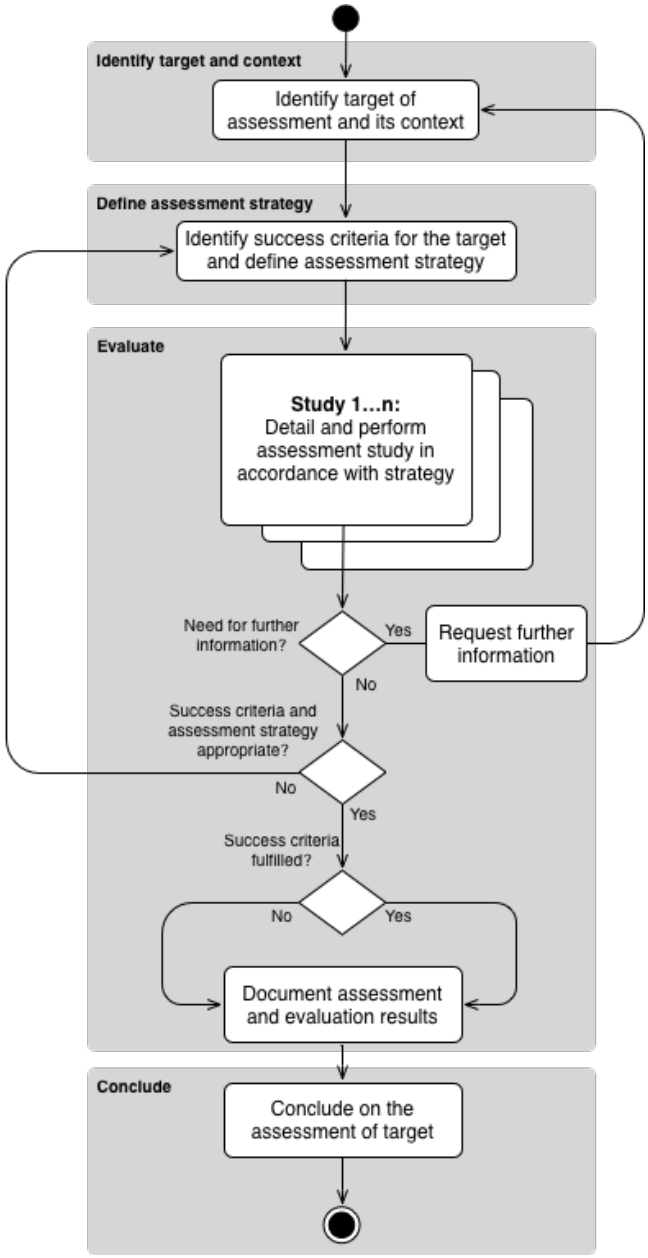


Figure 1 - The assessment process

3.1. The process

The assessment process proposed in Figure 1 is inspired by [21]. The first step is to identify the target of assessment and its context in order to facilitate establishing an effective assessment strategy. Depending on the nature of the target and the context for which it will be applied, the assessor defines the set of criteria that facilitates the evaluation of the target together with a strategy for performing the evaluation. In this context, the set of criteria is the claims and evidences extracted from the documentation or that the assessor defines himself, based upon the identified target and scoping of the investigation.

For each claim and evidence associated with the target that is within the scope of the investigation, the assessor must choose the best way of evaluating these. Typically, claims and evidences are extracted from the safety documentation and assessed, or the assessor has its own claims it needs investigated and evaluated. Independent of claims and evidences being extracted or self-defined, some criteria associated with them and their evaluation facilitates the assessor in deciding whether to accept, reject, or request more information. According to McGrath [22], there are eight common strategies for evaluation of predictions. However, there is no single evaluation method that provides results that are strong on generality, provides very precise measurements and at the same time is performed in environments that are very similar to reality. In the event that the assessor performs its own evaluation, he or she has to choose a set of evaluation methods and decide on the basis of the strengths and weaknesses of the different evaluation methods how different kinds of methods are best combined. In the event that extracted claims and evidences are assessed, the assessor needs to investigate whether an acceptable combination of methods have been applied by the licensee. The set of different kinds of evaluations the assessor apply in the investigative process represents the evaluation strategy. Depending on the results from assessing already performed evaluations and the availability of data to perform own studies appropriately, the assessor adapts the assessment process. The assessment process supports the need to adapt as new knowledge is acquired as can be seen from the back-loops in Figure 1.

Given that an evaluation study has the appropriate target data (see first decision in Figure 1) and that the assessment criteria and strategy is appropriate (see second decision in Figure 1), the evaluation results should be documented independent of the success criteria being fulfilled or not. The documentation of how the assessment was performed and the results of different evaluations is the grounds that supports the conclusion.

3.2. The language

An assessor must make decisions upon the basis of imperfect claims, arguments and evidence. Although it is difficult to make the human thinking process itself transparent, some transparency is gained if the main contributing factors leading to the decision is explicitly described. In order to distinguish between the elements of the safety argument and the assessment of these elements, we describe in the following sections a combined use of:

- An argumentation model – a model for expressing claims, arguments and evidence; and
- An assessment model – a model for expressing an assessor's plan for the evaluation of the content of an argumentation model as well as the results of evaluation.

3.2.1. The argumentation model

Figure 2 illustrates a small set of elements and how they may be used to capture a claim, its decomposition via relations into sub-claims that are supported by evidences. This small set of modelling elements is a simplification of notation offered in GSN [23] and CAE [7] notation. The intention of the simplification is to offer only those first-class citizens in the language that an assessor really needs and still offer expressional power. In order to reduce visual complexity, the argument logic (see "Argumentation model" in Figure 2) is illustrated separately from the definitions (see table "Definitions" in Figure 2) of the different elements. In Figure 2, colouring of argumentation elements

is used to distinguish between extracted and self-defined claims and evidences. It is up to the assessor to decide if the assessment will be performed purely on the basis of extracted claims and evidences, purely on self-defined claims and evidences, or a mix.

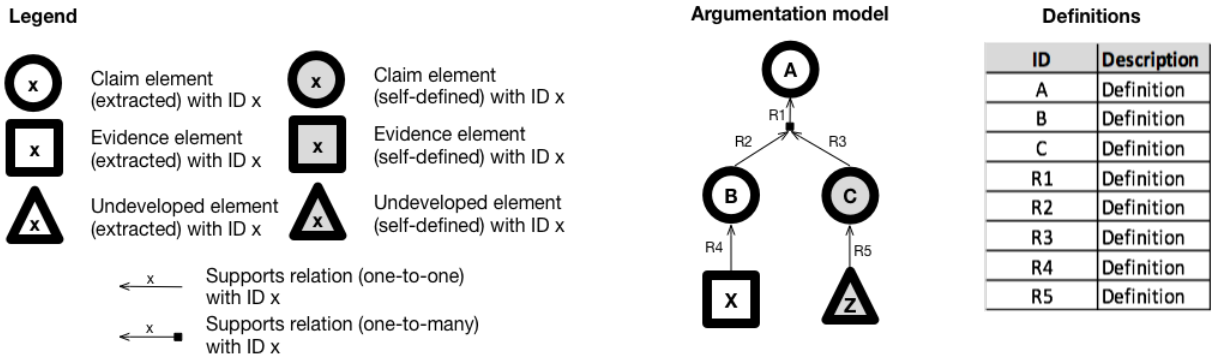


Figure 2 - An argumentation model

In Figure 2, claim "A" is decomposed into the sub-claims "B" and "C" where "B" is extracted from some safety documentation and "C is defined by the assessor. Whilst "R1" is a one-to-many relation, "R2" and "R3" are one-to-one relations. This way of modelling relations is chosen as once the "R1" decomposition into its parts is defined, then "R2" and "R3" relations should express a one-to-one relation between different parts in the "R1" relation and an associated element, here claim "B" and claim "C". A self-defined undeveloped element, as in "Z" in Figure 2, may represent the assessor intention to perform a particular kind of assessment suitable for investigating the support for claim "C". An extracted undeveloped element could represent a commonly accepted truth that is described in the safety documentation assessed but is not further evidenced as there is no need for that.

3.2.2. The assessment model

Figure 3 illustrates a small set of elements and how they may be used to capture an assessor's evaluation of an argument structure.

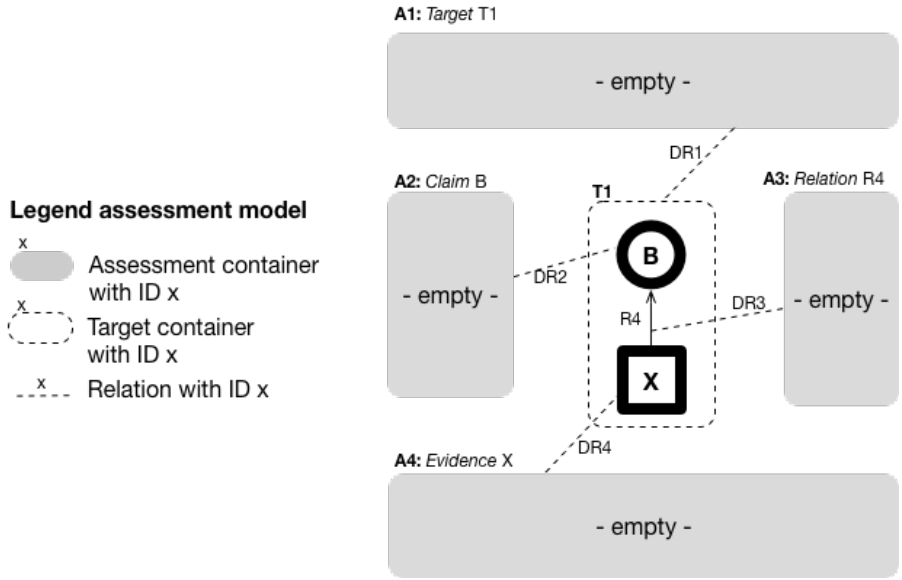


Figure 3 - A simple decision model

The assessment model consists of three kinds of elements as shown in the legend. The *assessment container* is intended to contain the assessment of an argumentation element (e.g. a claim, an evidence or a relation element in Figure 2) or a set of such elements. In Figure 3, the assessment containers are

deliberately denoted empty as a set of potential solutions is discussed further down. A *target container* is used to demark a set of argumentation elements that is intended to be assessed as a whole. A *relation* is used to connect an assessment (or more precisely the content of an assessment container) with the target of the assessment.

The intention with dedicated symbols for capturing the assessment of different elements of an argument structure is separation of concerns. The intent of the argumentation model is to express inferences from claims to evidences logically. The intent of the assessment model is to capture an assessor's evaluation of the different parts of an argumentation model.

In Figure 3, the results of the assessments would be described within the grey rectangles. However, depending the nature of the target of assessment, different kinds of assessment may be applied. Some approaches for assessing a claim, a relation or an evidence may be:

- To agree to a claim or accept an assumption as it expresses a commonly accepted truth or established practice.
- Perform a qualitative assessment according to an acceptable method.
- Reason on the basis of a statistical approach.
- Reason on the basis of logic or mathematical approach.

As different kinds of argument elements may be evaluated in different ways, results may not be easily combined. In addition, it can be challenging for an assessor to clearly define the transition from not being confident in the evidences to being confident or vice versa. Also, it is difficult to make sharp and crisp interferences from a claim down to the different evidences.

Although a probabilistic solution may be used for combining and aggregating individual assessment results, our main message is that the assessor at least should express the decisions from the assessment of individual argument elements explicitly. Then the assessor should decide on how to weight importance and express confidence to different parts of the argument structure and conclude on the basis of the sum of the evidence. In the exemplification of our strategy in Chapter 4, a probabilistic approach is discussed as support for concluding the assessment.

4. THE STRATEGY EXEMPLIFIED

In the following, our strategy for assessors is exemplified according to the four main steps of the process defined in Figure 1.

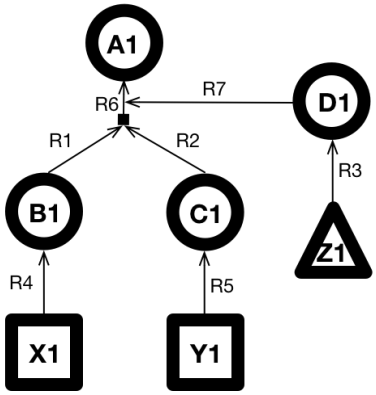
4.1. Step 1: Identify target and context

In the following example scenario, we assume that a third party is to assess a submittal of a DI&C system, where the submittal is a safety analysis report (SAR) on design of a DI&C system submitted to a nuclear regulator. With the SAR as the basis, information on potential claims and evidence for argumentation of independence of the system is extracted. The extracted information is organized into an argumentation structure. The argument structure is then assessed with respect to sufficiency in supporting the independence claim. The argument structure is presented somewhat abstract for illustrational purposes. However, it serves as a good example of what kind of safety argumentation contained in the documentation an assessor receives.

*…[**Claim A1 unfolded**]…because Claim B1 and Claim C1 [R1 & R2]. Claim B1 and Claim C1 together are equivalent with Claim A1 [**Claim D1 unfolded**] because we know [R3]…[**Undeveloped element Z1 unfolded**]… … Text with no safety argument relevance … … [**Claim B1 unfolded**]… because of [R4] Evidence X1. …[**Claim C1 unfolded**]… because of [R5] Evidence Y1.*

On the basis of the above textually described safety argumentation, an assessor may model it with the notation presented earlier into the argumentation structure in Figure 4. Each part of the argumentation structure and their relations is visualised. The left part of the figure only illustrates the organisation of argument elements and identifies them. The table on the right documents the content of identified argument elements.

In Figure 4, the relations "R6" and "R7" are added to the model in order to capture the one-to-many relation expressed in the textual description but not clearly identifies with separate ID´s.



ID	Description
A1	...[Claim A1 unfolded] ... because Claim B1 and Claim C1 [R1 & R2]
B1	...[Claim B1 unfolded] ... because of [R4]
C1	...[Claim C1 unfolded] ... because of [R5]
D1	Claim B1 and Claim C1 togheter are equivalent with Claim A1 [Claim D1 unfolded] ... because we know [R3]
X1	Evidence X1 ...
Y1	Evidence Y...
Z1	[Undeveloped element Z1 unfolded]... Text with no safety argument relevance ...
R1	... because Claim B1 and Claim C1 [R1 & R2]
R2	... because Claim B1 and Claim C1 [R1 & R2]
R3	...we know [R3] ...
R4	...because of [R4] Evicence X1
R5	...because of [R5] Evidcence Y1
R6	R6 express combination of B1 through R1 and C1 through R2 as support for A1
R7	R7 express that D1 supports ....

Figure 4 – A model of the inferences in the example safety document

4.2. Step 2: Define assessment strategy

In Figure 5, we assume that the assessor only explores the path from claim "A1" to evidence "X1" and leave the remaining two paths from "A1" to "Y1" and "A1" to "Z1" unexplored. One reason for this choice may be that the path "A1" to evidence "X1" is chosen in out of a number of trace-following investigations, or that the unexplored paths in the opinion of the assessor are not critical for safety but the chosen one is. Independent of the reasoning behind the selection, in order to gain transparency, the assessor should justify the selection.

In Figure 5, we have omitted visualising the assessment relations ID´s in order to reduce visual complexity. The assessment containers "Asm-1" to "Asm-6" addresses each of the argumentation elements in the path from claim "A1" to evidence "X1". The assessment "Asm-7" contains the assessor justification of not exploring the target named "T1", which contains the elements not explored. For visualisation purposes, the elements contained within "T1" is faded with a grey tone in order to indicate that these are unexplored.

A decision scale and a confidence scale inspired by [17] is used in Figure 5. An assessor´s decision is expressed by the colouring of the assessment containers according to the categories shown in the legend to the left in in Figure 5. An assessor´s confidence is expressed within an assessment container textually according to the following categories: Certain; High confidence; Medium Confidence; Low confidence; Lack of confidence. We assume in Figure 5 that the assessor expresses its decision on all elements and leave the confidence evaluation underspecified besides from in "Asm-5".

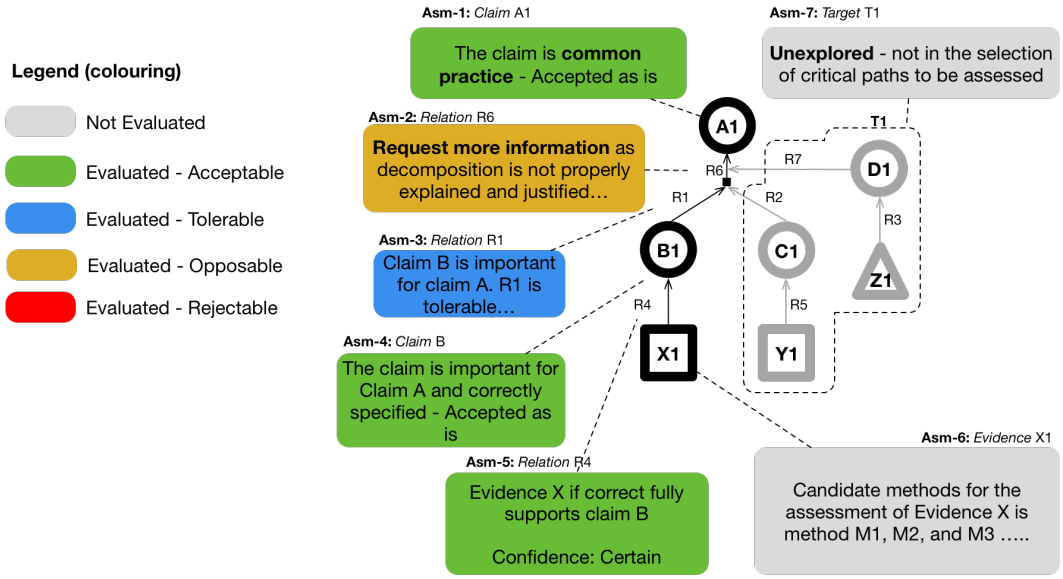


Figure 5 – Planning the assessment of the path Claim A1 to Evidence X1

4.3. Step 3: Evaluate

As stated earlier, depending on the nature of the target of assessment there are several potential ways of performing the assessment. In Figure 5, the assessor indicated in "Asm-6" three different methods that can be applied in order to assess evidence "X1". In Figure 6, we illustrate three different scenarios where evidence "X1" is evaluated with different methods and how the assessor may express the results of the assessment. The point to be made is that individual assessment results may not be easily combined into an assessment of the whole argument structure. Instead, we promote expressing the results of the assessment of elements (here by colouring) as support for the assessor in identifying visually those parts of the argument that are either acceptable, tolerable, opposable or rejectable decided on the basis of the level of confidence gained from assessing. In Figure 6, the different kinds of assessment are all denoted as acceptable, represented by the assessment container being coloured green. However, confidence is expressed differently in order to exemplify how different kinds of assessment can provide different kinds of confidence.

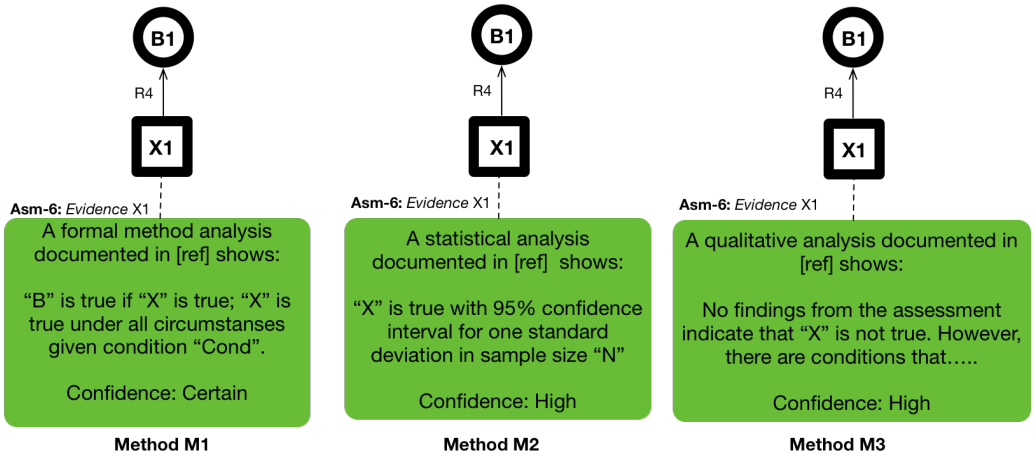


Figure 6 – Examples of documenting results from different kinds of assessments of an Evidence X1

4.4. Step 4: Conclude

Depending on the target of assessment and the assessor, it is expected that different approaches may be applied in how the assessor aggregate assessment results. Two very different approaches may be:

- Deterministic reasoning approach where the assessor gives more or less equal importance to separate parts of the argument and conclude with the proposal being acceptable only if every claim is acceptably substantiated.
- Probabilistic reasoning approach where the assessor weights separate parts of the argument differently and conclude with the proposal being acceptable or not although presented with imperfect claims and evidences.

We assume that the assessor through a combination of extracting and specifying self-defined claims and evidences has performed investigation of a set of claims. Furthermore, we assume the assessor has chosen a probabilistic reasoning approach to aggregate results. In Figure 7, the scope of our exemplification is shown, consisting of the "A1" claim discussed in the previous sections. Given adequate evidence of the "A1" claim, we assume that the assessor will find it acceptably supported. However, there are many options for how an assessor may apply probabilistic reasoning.

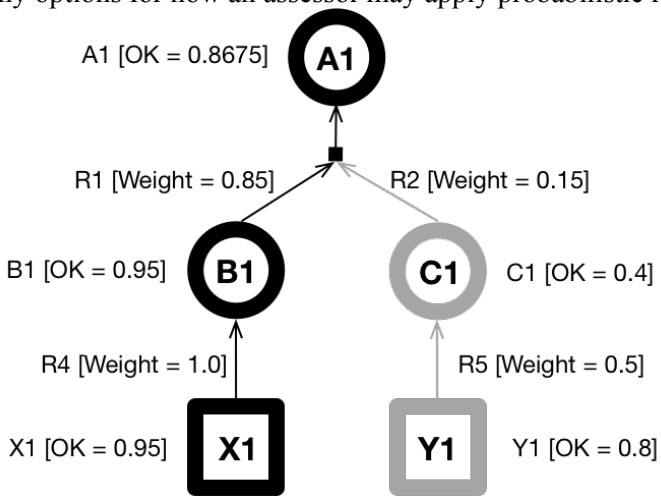


Figure 7 – Argumentation model as weighted graph example

One option is to apply the argumentation model as an influence diagram. In this case, each evidence can be assigned a belief function or a score, and each arc could be assigned a weight. This will allow to make an aggregated score or belief of the top node "A1". In Figure 7, the argumentation model in Figure 5 is presented as a simple weighted directed graph. We assume here that the weights on the relations and the values on the nodes are derived by the assessor on the basis of the assessments. The assessor uses the score for "A1" to support the conclusion on whether to accept or not.

Another option for the assessor is to apply a probabilistic approach using Bayesian Belief Network (BBN) to justify the conclusion. A BBN is a connected and directed graph, consisting of a set of nodes and a set of directed arcs between them. To each node there is associated uncertain variables, where the uncertainty is expressed by a probability density. The probability density expresses the confidence to the various variable outcomes and depends conditionally on the status of the "parent" nodes at the incoming edges. A node may be of various types, e.g. representing a continuous variable, or representing discrete states. The discrete states may be represented by e.g. Boolean, interval based, numbered or labelled states. The theoretical basis for making inference in BBN is based on Bayes rule, and a detailed description of the BBN methodology may be found in books by e.g., Jensen [18] and Pearl [19]. Today a large number of tools support the application of BBNs. A detailed description, with a simple example can be found in [20]. As discussed in [18] and [20], one advantage of applying the BBN approach is that one can divide the BBN modelling into two distinct parts; the modelling of the nodes and its connections (the network structure) and the modelling of the dependencies (the conditional probabilities assigned to each arc). These two steps are followed by calculations, which are based on the rules for conditional probability calculations. In our example, the first step would be to transfer the argumentation model into a BBN. This can be done by swapping the direction of all arcs in Figure 7. The node "A1" thereby becomes the parent of the children nodes "B1" and "C1", and for example, "X1" becomes the child of "B1". The next step is to decide the joint probability distribution

between a parent and a child. As each child here has only one parent, this is a simple conditional function or probability table. To simplify this, and to support consistency through a BBN it is often recommended to use templates [24]. By inserting the score on each evidence, these results will now be aggregated up the network using the Bayesian rule.

5. FURTHER WORK

This paper presents our initial proposal for supporting assessors in developing their assessment strategy and perform assessments. As a consequence, the focus has been on describing the strategy, and not to go into details. One such detail is the consequences of an influence or probabilistic model instead of the other options(s). One concern is that each option makes use of different sets of information such as weight, and each option interpret the information in a different way. In addition will the use of a probabilistic method as BBN require that we also adhere to the rules of Bayesian statistics. These are aspects that should be addressed in the further work.

A simple prototype tool named Instruct is developed at the OECD Halden Reactor Project that supports supervised extraction of claims, arguments and evidence from safety documentation (e.g. pdf files) and modelling of the extracted argument structure. Future work may be to offer more support on the assessment of extracted arguments as presented in this paper. Further work also includes providing a clear definition of the syntax and semantics of the language described. In addition, the graphical parts of the language need to be developed in accordance with state-of-the-art knowledge on visualisation for effective human comprehension. Most importantly, empirical evaluations with assessors are needed in order to evaluate its practical application and compare its use with other approaches.

6. CONCLUSION

As an assessor is likely to adapt its investigative process as new knowledge is acquired about the target, we propose a process and a language supporting the assessor in defining its strategy and collect results in a flexible manner. Furthermore, as the assessor may combine different approaches in the investigation of some system and application context, individual assessment results may not be easily combined into a well-defined conclusion of the whole. Judgement is needed, and the assessor combines different evidences on the basis of experience. For that reason, we describe a solution where the assessor explicitly expresses its evaluation of individual assessments in order to facilitate concluding. Furthermore, potential probabilistic reasoning approaches as support for an assessor when concluding its assessment was briefly discussed.

In the paper, our proposed strategy for assessors is only briefly described. The strategy needs to be further detailed and also needs empirical evaluation.

Acknowledgements

This work has been conducted within the OECD Halden Reactor Project, Institute for Energy Technology, Halden, Norway. Reports issued from the Halden Reactor Project, identified with "HWR" in the references are available upon request.

References

[1] A. A Hauge, P. Karpati and V. Katta, "*Summary of the 2014 Expert Workshop on Safety Demonstration and Justification of Digital Instrumentation and Control Systems in Nuclear Power Plants*", OECD Halden Reactor Project, HWR-1113, Halden, Norway, 2014.

[2] P. Karpati, V. Katta and C. Raspotnig, "*Expert Workshop on DI&C Safety Assurance with Special Focus on Experiences with Assurance Cases*", OECD Halden Reactor Project, HWR-1220, Halden, Norway, 2017.

[3] P. Karpati, A. A. Hauge, V. Katta and C. Raspotnig, "*Safety Demonstration and Justification of DI&C Systems for NPPs – Elicitation Interviews with regulators*", OECD Halden Reactor Project, HWR-1112, Halden, Norway, 2014.
[4] P. Karpati, K.C. Attwood, S. Nair, V. Katta and C. Raspotnig, "*Extracting the Safety Argumentation from an Interim Safety Demonstration – A Case Study from the Nuclear Field (Part 1: Argument Comprehension)*", OECD Halden Reactor Project, HWR-1149, Halden, Norway, 2016.
[5] P. Karpati, K. C. Attwood, S. Nair, V. Katta and C. Raspotnig, "*Improving Safety Arguments for Better Comprehension – Lessons Learned from a Case Study in the Nuclear Field*", OECD Halden Reactor Project, HWR-1193, Halden, Norway, 2016.
[6] SSM, *"Licensing of Safety Critical Software for Nuclear Reactors, Common Position of Seven European Nuclear Regulators and Authorised Technical Support Organisations – Revision 13"*, Swedish Radiation Safety Authority, Report number 2013:08, 2013, Stockholm, Sweden
[7] GSN Community Standard. Version 1.0. (2011). Accessed March 2018, available: http://www.goalstructuringnotation.info/
[8] L. Emmet, G. Cleland, "*Graphical Notations, Narratives and Persuasion: a Pliant Systems Approach to Hypertext Tool Design*", In Proceedings of ACM Hypertext 2002 (HT'02), June 11-15, College Park, Maryland, USA, 2002.
[9] Interim Defence Standard 00-56 Issue 5", UK MOD, 2014.
[10] CENELEC, "*EN-50129 Railway Applications - Communication, signalling and processing systems — Safety related electronic systems for signalling*", 2003.
[11] EUROCONTROL, "*Safety Case Development Manual*", 2006.
[12] S. Toulmin, "*The Uses of Argument*", Cambridge, University Press, 1958.
[13] J. Górski, "*Trust case—a case for trustworthiness of IT infrastructures. Cyberspace Security and Defense: Research Issues*", pages 125-141, Springer Netherlands, 2005.
[14] W. S. Greenwell, J. C. Knight, C. M. Holloway, and J. J. Pease, "*A Taxonomy of Fallacies in System Safety Arguments*", Proceedings of the 24th International System Safety Conference, Albuquerque, New Mexico, USA, 2006.
[15] R. Hawkins, T. Kelly, J. Knight and P. Graydon, "*A new approach to creating clear safety arguments*", In Advances in Systems Safety, pages 3-23, 2011.
[16] S. Nair, N. Walkinshaw, T. Kelly and J. L. d. l. Vara, "*An Evidential Reasoning Approach for Assessing Confidence in Safety Evidence*", Simula Research Laboratory, Technical Report, 2014.
[17] J. Górski, "*Trust-IT – a framework for trust cases, Workshop on Assurance Cases for Security - The Metrics Challenge*", DSN 2007 The 37th Annual IEEE/IFIP Intern. Conf. on Dependable Systems and Networks, Edinburgh, UK, June, 2007.
[18] F. Jensen, "*An Introduction to Bayesian Network*", UCL Press, 1996.
[19] J. Pearl, "*Probabilistic Reasoning in Intelligent Systems: Networks for Plausible Inference*", Morgan Kaufman, 1988.
[20] B. A. Gran, "*Use of Bayesian Belief Networks when Combining Disparate Sources of Information in the Safety Assessment of Software Based Systems*", in International Journal of Systems Science, 33 (6): 529-542, 2002.
[21] T. P. Kelly, "*Reviewing assurance arguments – a step-by-step approach*", in Proc. Workshop on Assurance Cases for Security –The Metrics Challenge, 37th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, 2007.
[22] J. E. McGrath, "*Groups: interaction and performance*", Prentice-Hall, 1984.
[23] T. P. Kelly, "*Arguing Safety – A Systematic Approach to Safety Case Management*", DPhil Thesis, YCST-99-05, Department of Computer Science, University of York, UK, 1998
[24] J. E. Simensen, G. Gerst, B. A. Gran, J. Märtz, H. Miedl; "*Establishing the Correlation Between Complexity and a Reliability Metric for Software Digital I&C-Systems*" In LNCS 5775, Springer Berlin-Heidelberg, pages 55-66, 2009.