Belief functions for safety arguments confidence estimation : A comparative study

Yassir Idmessaoud, Didier Dubois, Jérémie Guiochet

Belief functions for safety arguments confidence estimation : A comparative study

Yassir Idmessaoud[1], Didier Dubois[2], and Jérémie Guiochet[1]

[1] LAAS-CNRS, University of Toulouse, France
[2] IRIT, University of Toulouse, France

Abstract. Structured safety arguments are widely applied in critical systems to demonstrate their safety and other attributes. Graphical formalisms such as Goal Structuring Notation (GSN) are used to represent these argument structures. However, they do not take into account the uncertainty that may exist in parts of these arguments. To address this issue, several frameworks for confidence assessment have been proposed. In this paper, a comparative study is carried out on three approaches based on Dempster-Shafer theory. We extract and compare the implicit logic at work in these works, and show that, to some extent, these current approaches fail to provide a consistent relationship between the informal statement of arguments, their logical model and the use of belief functions. We also propose recommendations to improve this consistency.

Keywords: Confidence Assessment · Goal Structuring Notation (GSN) · Dempster–Shafer theory (DST) · Evidence fusion · Safety cases.

1 Introduction

The deployment of autonomous systems in the highly uncertain human environment raises the issue of safety. Argument structures are widely used to evaluate and prove the safety of these systems. They are a clearly represented collection of rational pieces of evidence like test or simulation results, expert judgments, analysis reports, etc. They aim to demonstrate that a certain property of the system is satisfied. Many studies and standards define safety arguments as "Safety cases" (e.g, in the automotive [17] or railway [12] industries), but it is now extended to more general domains like dependability, assurance or trust cases. These cases are presented in the form of texts, tables or, more interestingly, graphically. Using graphical tools to represent arguments structures is more relevant because graphs are simpler to review, offer a clear overlook, help to understand the connection between pieces of evidence, and moreover they are easier to use and manage. Formalisms such as Goal Structuring Notation (GSN) [19] and Claims-Arguments-Evidence (CAE) [2] are commonly used in this field. However, even with all these benefits, these tools do not take into account the uncertainties pervading this sort of arguments. Especially since autonomous systems, and critical systems in general, are becoming much more complex, they are affected by many sources of uncertainty like any decision support system, AI based system and

the like. As a response to this issue many research projects are conducted to find a solution.

Several works have proposed methods based on Bayesian Networks (BN) to model uncertainty in a safety argument [15, 6, 14]. Nevertheless, these approaches have a major impediment, which is the need for *data*. As a matter of fact, using BN requires statistical information that is often not available. Moreover, the use of subjective probabilities is questionable in the presence of partial ignorance. Dempster-Shafer theory (DST) (aka Theory of Evidence) was developed to address the issue of imprecise evidence [21]. It represents a form of generalized probability theory where probability masses are assigned to sets of possible values, instead of singletons. In some works on safety argumentation, aggregation rules stemming from DST are used to merge confidence degrees in pieces of evidence (represented by mass functions) and calculate an overall mass function, in order to estimate an overall confidence in the top statement of a safety argument (e.g.,"the system is acceptably safe"). In our problem, the connection between pieces of evidence is represented by various types of arguments. In the literature, they strongly influence the choice of an aggregation rule. However, no real consensus emerges in current research works to relate argument types, their logical modeling, and aggregation rules based on DST.

In this paper, three approaches to uncertain safety cases using DST are compared as to the different definitions of types of arguments they propose and we review and discuss the aggregation methods they use. Section 2 presents the background on safety cases and introduces the existing selected approaches. Section 3 extracts the formal definitions of arguments from the selected articles [1, 4, 24]. Then, we compare and analyze the aggregation rules used to compute belief degrees of the top statement of an argument. Section 4 suggests the existence of two basic types of arguments and proposes a rigorous methodology.

2 Baseline and related work

This section introduces a safety case formalism (GSN), some works on confidence quantification, and basic concepts of DST.

2.1 Background

Safety arguments or safety cases can be defined in multiple ways. In fact, the definition may vary slightly according to the field where it is used. For instance, in the automotive industry [17], it is defined as : *argument that functional safety is achieved for items, or elements, and satisfied by evidence compiled from work products of activities during development.* This concept has been generalized in the OMG (Object Management Group) standardized Assurance Case Metamodel [5] and an instance of it is the goal structuring notation (GSN), which is commonly used to represent safety cases [19]. As presented in Figure 1, it includes nine main elements. It breaks down the conclusion called a *Goal* (following a given *Context* and in accordance with a specific *Strategy*) into *Sub-goals* and

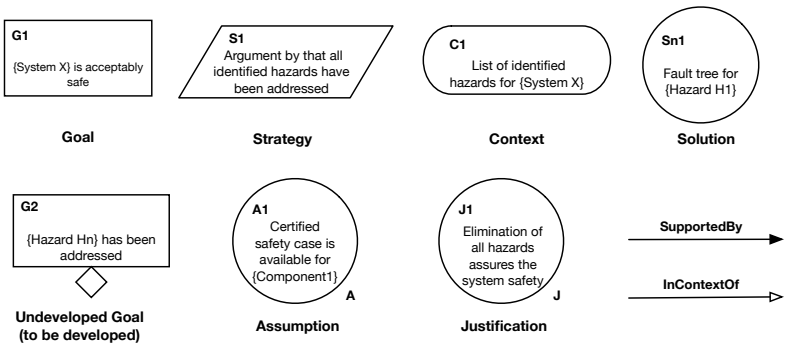

Fig. 1. GSN main components.

supports each of them by evidence items called *Solutions*. The choice of strategies and sub-goals is supported by the use of so-called *Justifications*. Figure 2 presents an example of a GSN pattern.

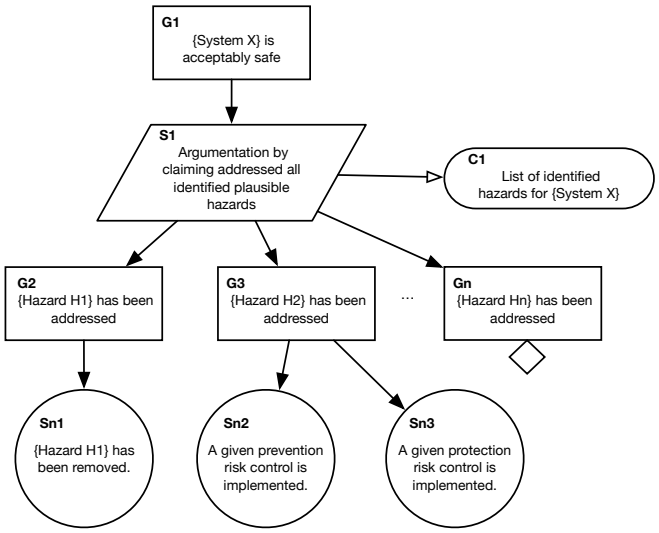

Fig. 2. GSN example adapted from Hazard Avoidance Pattern [20].

GSN is categorised as a qualitative method to justify safety. However, in this example, many uncertainties may exist. For instance, what is the uncertainty linked to the element "C1: List of identified hazards" or what is the confidence in the solutions Sn (also called pieces of evidence). In order to estimate the confidence in the top goal $G1$, all these uncertainties should be assessed. Some

quantitative approaches have been proposed to assess uncertainties in such arguments [13]. In[15, 6, 14], Belief Bayesian Networks (BBN) are used to assess confidence in safety case structure and pieces of evidence. They measure confidence by computing probabilities from evidence to conclusion. Due to the huge amount of data required to apply BBN, other works based on subjective logic [18, 29] or on DST are proposed to define and estimate uncertainties. These works are presented in Section 2.2.

Dempster-Shafer Theory offers a powerful setting to combine pieces of evidence. A mass function, or basic belief assignment (BBA), assigns probabilities over the power set of the universe of possibilities (Ω), known as the *frame of discernment*. Formally, a mass function $m^{\Omega} : 2^{\Omega} \rightarrow [0, 1]$ is such that $\sum_{E \subseteq \Omega} m(E) = 1$, and $m(\emptyset) = 0$. Any subset E of Ω such as $m(E) > 0$ is called a focal set of m. Mass assignment induces the concept of belief function ($bel : 2^{\Omega} \rightarrow [0, 1]$). It represents the summation of all the masses supporting the same statement and is defined by : $bel(A) = \sum_{E \subseteq \Omega, E \neq \emptyset} m(E)$. Belief in the denial or uncertainty of the statement A are respectively represented by : $disb(A) = bel(\neg A)$ and $uncer(A) = 1 - bel(A) - disb(A)$.

In our case, we use propositional variables for which the frame of discernment has two states: $\Omega = \{True(T), False(F)\}$. As a consequence, in such frames, mass function and belief function for T and F are equal. For example, consider a statement A saying that a clock provides the right time. The mass function m^{Ω} such that $m^{\Omega}(\{T\}) = 0.5$, $m^{\Omega}(\{F\}) = 0.2$ and $m^{\Omega}(\{T, F\}) = 0.3$ quantifies respectively the degrees of belief (0.5), of disbelief (0.2) and of uncertainty (0.3) in A. In this case, due to the Boolean form of the frame of discernment, we have $bel(A) = m^{\Omega}(\{T\})$, $disb(A) = m^{\Omega}(\{F\}) = 0.2$ and $uncer(A) = m^{\Omega}(\{T, F\}) = 1 - m^{\Omega}(\{T\}) - m^{\Omega}(\{F\})$. They represent respectively our belief that the time read is correct, not correct, and the probability that we don't know the time by reading the watch (Tautology).

Another important tool from DST is the Dempster rule of combination. It is used to merge various pieces of evidence coming from independent sources of information, and is represented by mass functions $m_i, i = 1, \ldots, n$. It proceeds in two steps. For two mass functions:

1. a conjunction of random sets: $m_{\cap} = m_1 \otimes m_2$ such that $m_{\cap}(A) = \sum_{E_1, E_2 : E_1 \cap E_2 = A} m_1(E_1) \cdot m_2(E_2)$;
2. a renormalization step if $m_{\cap}(\emptyset) > 0$: $m(A) = m_{\cap}(A)/(1 - m_{\cap}(\emptyset))$. The value $m_{\cap}(\emptyset)$ represents the degree of conflict between m_1 and m_2.

This combination rule is commutative and associative.

2.2 Some DST-based approaches to safety cases

In this subsection, we discuss three uncertainty management methods in safety cases proposed in the literature. All such methods are DST-based. Two of them use Goal Structuring Notation (GSN) for structuring arguments. Our objective is to extract, in each paper, definitions of argument types, and to evaluate their

consistency with the proposed aggregation rule that computes degrees of belief of conclusions.

Cyra and Gorski [4] present an argument model called VAA inspired by Toulmin [23] to graphically represent all pieces of evidence that support a conclusion (e.g., "The system is safe"). It proposes a method that, in a first step, transforms qualitative expert opinions expressed in natural language, about pieces of evidence (forming premises), into belief and plausibility functions, using [18]. In a second step, the authors define five argument patterns and associate to each of them an appropriate belief aggregation rule. These rules use as inputs the values obtained at the first step, to calculate the overall confidence in the conclusion.

Anaheed et al. [1] define four basic types of arguments. Each argument type is composed of at least two premises that support a conclusion. Premises are assessed by two parameters (sufficiency and insufficiency) and every argument type is associated to an aggregation rule. In this paper, they propose an algorithm based on a bottom-up approach that computes the degree of confidence in each premise to calculate the overall confidence in the system.

Wang et al. [25, 24, 26, 27] propose a confidence assessment method by converting qualitative expert opinions, on their confidence in pieces of evidence appearing in a GSN, into mass functions. These values are then merged by Dempster rule of combination to obtain the overall confidence in the studied system. The paper also defines two parameters to assess confidence: *Trustworthiness* to quantify confidence in the evidence; *Appropriateness* to quantify the confidence in the claim that the evidence supports the conclusion.

3 Comparative study

In this section, we introduce a framework for confidence assessment. Then, we compare argument types given in the studied papers and the propagation rules used to calculate the overall confidence.

3.1 General framework for BF-based confidence estimation

In order to estimate the overall confidence of the argument structure, our main issue is how to propagate the quantitative values coming from the confidence in premises in accordance with the characteristics of its structure. In this regard, it is important to propose a general method to compute degrees of belief in conclusions of safety cases. Most works omit to provide this general method putting together logic and belief functions. Such a methodology was described more than 30 years ago [3] and is recalled here.

The first step is to define the nature of the relationship between premises in their support of the conclusion, known as *argument types*. These types should be firstly expressed informally in a natural language (e.g, if premises P_1 and P_2 are true, then the conclusion (C) is true) because it is more understandable for the human expert. Then this verbal relation should be transformed into a formal logical sentence (e.g., $P_1 \wedge P_2 \Rightarrow C$). The importance of these definitions lies in

the fact that they significantly affect the logical expression of the links between pieces of evidence and the conclusion. The second step is mass assignment. This task consists in defining masses assigned to the focal sets deduced from the logical expressions obtained from argument definitions.

Consider a set of well-formed formulas $K = \{\phi_1, \dots \phi_n\}$ in propositional logic, and a formula C such that $K \vdash C$. Assume each formula ϕ_i is a piece of evidence that comes from a specific source independent of other ones. Uncertainty about the validity of each formula can be represented by a mass function m_i assigning some probabilities to $\phi_i, \neg\phi_i$ and the tautology $\top$. Take for example the case of a simple premise P and a rule in the form of an equivalence $P \equiv C$. One mass function will be assigned to the premise P in the form of three values $m_1(P)$, $m_1(\neg P)$ and $m_1(\top)$ summing to 1, and another will be assigned to the rule $(m_2(P \equiv C) + m_2(\top) = 1)$.

The third step is to choose the appropriate aggregation rule. This rule will be used to calculate the belief in the top goal (conclusion) based on beliefs about premises and rules. Extending classical logic inference to this uncertain environment can be done by means of Dempster rule of combination [3], first computing an overall mass function, $m = m_1 \otimes \cdots \otimes m_n$ and then computing the degree of belief in the conclusion C as $Bel(C) = \sum_{\phi_i \vdash C} m(\phi_i)$. There are also several variants of this combination rule that could be used in evidence fusion. However, each method obeys certain assumptions and describes some kind of situation. That is why it is needed to make sure that every definition resulting from the first step verifies the assumptions and each fusion rule fits with the given situation. Here, pieces of evidence and rules are supposed to come from independent sources. If this assumption is not satisfied, idempotent combination rules can be used as discussed in [7, 8].

The complete process includes an additional preliminary step, which consists in transforming expert opinions (qualitative values) expressed in natural language (*safe, little safe, uncertain*, etc.) into a numerical format that can be computed with (i.e. mass, belief or plausibility functions). This could be done in [4] using the triangle method of Josang [18]. This is needed to compute the belief in the conclusion. The choice of this transformation has a profound impact on the results. However this aspect of the evaluation process will not be addressed in this paper.

3.2 Definition of argument types

The concept of *argument type* pertains to the logical relationship between the premises and the conclusion. In other words, it answers the question : In which format do the premises support the conclusion ? The terminology is not uniform. For instance, in [4] this relation is named a *warrant*, in [29] it is called an *affection factor* and in [24] it is named *appropriateness*. Moreover, most papers only give an informal definition.

Table 1 presents formal definitions of argument types that we infer from the reviewed papers. We notice from the formal definitions given in Table 1 that the premises are related to the conclusion by either an equivalence or an

Table 1. Formal definitions of arguments. Note that argument types 4 and 5 are logically equivalent ($\wedge_{i=1}^{n}[p_i \Rightarrow C] \equiv [\vee_{i=1}^{n} p_i] \Rightarrow C$).

	Formal definition	Terminology for argument types
Type 1	$(\wedge_{i=1}^{n} p_i) \equiv C$	NSC-Arg [4], Consensus [29], Full complementary [24]
Type 2	$\wedge_{i=1}^{n}(p_i \equiv C)$	Disparate [24]
Type 3	$(\wedge_{i=1}^{n} p_i) \Rightarrow C$	SC-Arg, C-Arg [4], Conjunctive argument [29]
Type 4	$\wedge_{i=1}^{n}(p_i \Rightarrow C)$	A-Arg [4], Alternative argument [1]
Type 5	$(\vee_{i=1}^{n} p_i) \Rightarrow C$	Disjunctive argument [29], Full redundant [24]
Type 6	$\begin{cases} \wedge_{i=1}^{n}(p_i \equiv C) \\ (\wedge_{i=1}^{n} p_i) \equiv C \end{cases}$	AC-Arg [4], Complementary [24], Containment, Overlapping arguments [1]
Type 7	$\begin{cases} \wedge_{i=1}^{n}(p_i \equiv C) \\ (\vee_{i=1}^{n} p_i) \equiv C \end{cases}$	R-Arg [24]

implication connective. The choice could be justified by the intuitive perception of the relation between premises and conclusion (e.g. "The system is safe" is supported by "Hazard H1 has been addressed").

As already noticed, there are two types of arguments, one using implication, the other using equivalence. Using equivalence assumes that there is a symmetry between the conclusion and premises. Consider a small safety case where the statement "The system is safe" is supported by the premise "All tests are conclusive". Using equivalence means, on the one hand, that the system is safe because we are confident in our tests; on the other hand, that the actual safety of the system can only be ensured by the success of the tests, which is not necessary true. In addition to this, the use of equivalence generates cases for the denial of the conclusion (disbelief) which appears in the calculation as a conjunction of one or several premises with the negation of the conclusion ($\neg C$). For instance, consider the case of a conclusion (C) supported by a single premise (p) with an equivalence relation between them (Type 1 or 2). Combining the masses of (p) and the rule ($p \equiv C$) with DS combination rule reveals two cases where the conclusion is not satisfied ($\neg p \wedge \neg C$) and ($p \wedge \neg C$). Using implication can only indicates that, due to the tests, the system is safe; it cannot prove that it is faulty. Choosing between equivalence or implication could also be justified by the purpose of the safety case. Generally speaking, a safety case is used to demonstrate that a system is acceptably safe. Its purpose is to provide a structured argument in order to certify a critical system, and not to present statements that it could be faulty, i.e. that there is disbelief about safety greater than 0. This is actually guaranteed when using only implication. In contrast, if the goal is to use the safety case at the debugging phase, i.e. to consider that disbelief in the conclusion may be not null, the equivalence may be an appropriate choice between premises and conclusion. Since we are interested, in this study, in the certification aspect of safety cases, the remainder of the paper will be focusing on argument types modelled by implication from tests to a statement of safety.

We also notice from Table 1 that the premises are linked with each other by AND, OR logic gates or by a combination of the two. It depends on whether,

for instance, tests justifying a conclusion are alternative or complementary. For example, type 3 represents the situation where the conjunction of all premises is needed to support the conclusion. In the contrary, type 4 represents the case of separate rules where each premise alone can support the whole conclusion.

3.3 Mass assignment

As seen in previous sections, masses are allocated to propositions of interest. Apart from the assignment of masses to logical expressions resulting from the definition types arguments (called *appropriateness* in [24]), masses are also assigned to premises to assess their degree of confidence. This evaluation is used under the name *trustworthiness* in [24] and *affection factor* in [29]. Normally, mass functions assigned to premises (P_i) have two parameters : belief (i.e. $m_p(P_i)$), disbelief (i.e. $m_p(\neg P_i)$) and the remainder is their uncertainty (i.e $m_p(\top) = 1 - m_p(P_i) - m_p(\neg P_i)$). In cases when the argument is an implication, not an equivalence, the disbelief in the premises will not affect the conclusion, and need not be taken into account in the uncertainty propagation. This remark reduces the number of useful focal sets and simplifies the calculation.

The choice of mass functions is a very important step in the assessment process. It has a huge impact on the form of the final result. We can either define several mass functions, one for each logical expression, to emphasize the fact that there are multiple independent sources of information. Or, one mass function is distributed over all the logical expressions to represent the situation where a single source supplies these pieces of information.

3.4 Belief Aggregation

As we saw in the previous Section 3.2 the informal definition of argument types is important for the belief assessment process. Since masses are also assigned to the logical formulas resulting from these definitions, many authors confuse logical and numerical aspects. But as explained in Section 3.1, the definition of argument type, especially the informal ones, conditions the choice of focal sets, on the one hand. On the other hand, it also affects the choice of the aggregation rule. For example, one may think of using the disjunctive consensus rule [10], if disjunction is expressed in the definition. In this section, we are interested in choosing aggregation rules based on Dempster-Shafer Theory and observing the effect of mass functions assignment on the degree belief of the conclusion.

DST offers many aggregation rules (Dempster rule, disjunctive consensus, Yager's rule, etc. see [21, 22] for surveys). However, we are going to focus on the methods used in the studied papers listed in Table 2. In general, Dempster combination rule computes the intersection of focal sets. If some focal sets from one source are inconsistent with some from another source, a renormalization must take place. It also assumes that sources are independent and reliable. Other combination rules that express a conjunction exist (e.g. Yager's [28] and Inagaki's rules [16]). For instance, Yager's rule uses a renormalization scheme different from the one of Dempster, reallocating mass of the empty set to the whole

frame. The disjunctive rule of combination is a union of random sets and does not need renormalization. Finally, the weighted average rule [21] is used to make a trade-off between conjunctive and disjunctive methods.

So long as the focal propositions involved in a safety case are not conflicting, there is no need to renormalize the resulting mass function, nor to use the disjunctive rule. By taking into account the definition of a safety case and the underying assumptions, applying Dempster rule (without normalization) to aggregate evidences is well adapted.

Table 2. Consistency between argument types and combination rules.

Authors	Argument types	Combination rules	Consistency
Cyra and Gorski [4]	NSC-Arg	DS rule	Yes
	SC-Arg	DS rule	Yes
	A-Arg	Yager's rule	No
	C-Arg	Weighted average	No
	AC-Arg	-	-
Anaheed et al. [1]	Alternative	DS rule	No
	Disjoint	Weighted average	No
	Containment	DS + Weighted average	No
	Overlapping		No
Wang et al. [24]	Disparate	DS rule	Yes
	Complementary		Yes
	Full complementary		Yes
	Redundant		Yes
	Full redundant		Yes

As can be seen from Table 2, several combination rules are proposed to obtain the overall confidence in the conclusion. But, none of the reviewed papers clearly justifies its choice of the applied method, nor does it lay bare the underlying independence assumptions. Also, some of the proposed expressions are not consistent with the verbal definition. For instance, in [1], the type *Alternative argument* is used when several independent premises support the conclusion. The formal definition induced is : $\wedge_{i=1}^{n}(p_i \Rightarrow C)$ (Table 1). However, the authors only consider the confidence in the premises (also called trustworthiness, in [24]), but they did not consider the confidence in the relation between them and the conclusion (the rule). A possible formula that takes into consideration this definition could be equation (1).

Consider the example of a conclusion (C) separately supported by two premises p_1 and p_2, which refers to an argument of type 4 in Table 1, i.e., $(p_1 \Rightarrow C) \wedge (p_2 \Rightarrow C)$. We develop this example below. For other argument types, the calculation follows the same method. The confidence assessment process is measured through two parameters. The confidence in premises is modelled by mass functions m_{p_i} and the confidence in the support of the the conclusion by each premise (the rules) is modelled by the mass function m_{r_i}. Then, we apply Dempster combi-

nation rule, presented earlier, to merge each premise with its appropriate rule (see Table 3), and secondly merge the two resulting mass functions m_i (see Table 4). Notice that we could also merge premises and rules separately first, then fuse partial conclusions together. The result will be the same because Dempster rule is associative and commutative.

Table 3. Combination of a premise with its rule

$m_1 = m_{p_1} \otimes m_{r_1}$	$m_{r_1}(p_1 \Rightarrow C)$	$m_{r_1}(\top)$
$m_{p_1}(p_1)$	$p_1 \wedge C$	p_1
$m_{p_1}(\top)$	$p_1 \Rightarrow C$	$\top$

In the example given in Table 3, the focal formula $p_1 \wedge C$ results from the conjunction between formulas p_1 and $p_1 \Rightarrow C$. Its mass is calculated by multiplying the masses values in the corresponding line and column. Since the frame of discernment (Ω) of elementary mass functions has two states, masses and belief functions of non-tautological inputs are equal. The calculation of the remaining masses follows the same logic. An example is given here, where, $bel^i_{\Rightarrow}(p_i \Rightarrow C)$ represents the degree of belief that the i^{th} premise supports the conclusion and $bel^i_p(p_i)$ represents the belief degree in the i^{th} premise. For instance, $m_1(p_1 \wedge C) = m_{p_1}(p_1) \times m_{r_1}(p_1 \Rightarrow C) = bel^1_p(p_1) \times bel^1_{\Rightarrow}(p_1 \Rightarrow C)$. Likewise the combination of m_1 and m_2, yields mass function m_{12} using Table 4.

Table 4. Combination of confidence in type 4 : $\wedge_{i=1}^{n}(p_i \Rightarrow C)$

$m_{12} = m_1 \otimes m_2$	$m_2(P_2 \wedge C)$	$m_2(P_2)$	$m_2(P2 \Rightarrow C)$	$m_2(\top)$
$m_1(P_1 \wedge C)$	$P_1 \wedge P_2 \wedge C$	$P_1 \wedge P_2 \wedge C$	$P_1 \wedge C$	$P_1 \wedge C$
$m_1(P_1)$	$P_1 \wedge P_2 \wedge C$	$P_1 \wedge P_2$	$P_1 \wedge (P_2 \Rightarrow C)$	P_1
$m_1(P_1 \Rightarrow C)$	$P_2 \wedge C$	$P_1 \wedge (P_2 \Rightarrow C)$	$(P_1 \vee P_2) \Rightarrow C$	$P_1 \Rightarrow C$
$m_1(\top)$	$P_2 \wedge C$	P_2	$P_2 \Rightarrow C$	$\top$

The calculation of the degree of belief $bel_c(C)$ in the conclusion for type 4 arguments is as follows (in Table 4 cells in gray identify ϕ's that imply C):

$$\begin{aligned}
bel_c^4(C) &= \sum_{\phi:\phi \ implies \ C} m_{12}(\phi) = m_{12}(P_1 \wedge P_2 \wedge C) + m_{12}(P_1 \wedge C) + m_{12}(P_2 \wedge C) \\
&= m_1(P_1 \wedge C) \sum_{\phi_2} m_2(\phi_2) + m_2(P_2 \wedge C) \sum_{\phi_1} m_1(\phi_1) - m_1(P_1 \wedge C) m_2(P_2 \wedge C) \\
&= m_1(P_1 \wedge C) + m_2(P_2 \wedge C) - m_1(P_1 \wedge C) m_2(P_2 \wedge C) \\
&= bel_p^1(P_1) bel_{\Rightarrow}^1(P_1 \Rightarrow C) + bel_p^2(P_2) bel_{\Rightarrow}^2(P_2 \Rightarrow C) \\
&\qquad - bel_p^1(P_1) bel_{\Rightarrow}^1(P_1 \Rightarrow C) bel_p^2(P_2) bel_{\Rightarrow}^2(P_2 \Rightarrow C) \\
&= 1 - [1 - bel_p^1(P_1) bel_{\Rightarrow}^1(P_1 \Rightarrow C)][1 - bel_p^2(P_2) bel_{\Rightarrow}^2(P_2 \Rightarrow C)]
\end{aligned}$$

In general, with n premises, the formula for type 4 arguments is as follows.

$$bel_c^4(C) = 1 - \prod_{i=1}^{n} [1 - bel_p^i(p_i) bel_{\Rightarrow}^i(p_i \Rightarrow C)] \tag{1}$$

Letting $bel_c^i(C) = bel_p^i(p_i) bel_{\Rightarrow}^i(p_i \Rightarrow C)$ be the degree of belief in C due to premise p_i, the expression in equation (1) is a many-valued disjunction connective aggregating the weights $bel_c^i(C)$. So it is enough that $bel_c^i(C) = 1$ for some p_i to get $bel_c(C) = 1$, which is in agreement with the argument type.

It is important to mention that in equation (1), a mass function $m_{\Rightarrow}$ was assigned to each rule $p_i \Rightarrow C$, assuming independence between them, according to type 4 in Table 1. In type 5 argument, we assign a single mass function $m_{\Rightarrow}$ to the complete rule with a disjunction of premises $([\vee_{i=1}^n p_i] \Rightarrow C)$. The formula resulting from this new mass assignment is given in (2). In general, the belief in the conclusion for type 5 arguments is as follows, using Dempster rule of combination:

$$bel_c^5(C) = bel_{\Rightarrow}([\vee_{i=1}^n p_i] \Rightarrow C)[1 - \prod_{i=1}^{n} (1 - bel_p^i(p_i))] \tag{2}$$

where $bel_{\Rightarrow}([\vee_{i=1}^n p_i] \Rightarrow C)$ is the belief that the disjunction of all premises support the conclusion. We stress again that in equation (1), we assign one mass function to each simple rule, while in (2), we assign a single mass to a composite rule. So, even though types 4 and 5 are logically equivalent $(\wedge_{i=1}^n [p_i \Rightarrow C]) \equiv ([\vee_{i=1}^n p_i] \Rightarrow C)$, because the assignment of masses is different in types 4 and 5, they produce different results for the belief calculation. The same combination pattern applies to arguments of type 3 in Table 1. It requires all premises be true to justify the conclusion, and a simple support mass is assigned to the implication $[\wedge_{i=1}^n p_i] \Rightarrow C$. It yields for type 3 arguments:

$$bel_c^3(C) = bel_{\Rightarrow}([\wedge_{i=1}^n p_i] \Rightarrow C) \prod_{i=1}^{n} bel_p^i(p_i) \tag{3}$$

In (2) a multivalued disjunction connective is applied to the degrees of belief in the premises while in (3), it is a multivalued conjunction.

In contrast, equation (4) below for type 4 arguments supposes a single mass function $m_{\Rightarrow}$ with masses distributed over the elementary rules $p_i \Rightarrow C$, assuming $\sum_{i=1}^{n} m_{\Rightarrow}(p_i \Rightarrow C) + m_{\Rightarrow}(\top) = 1$. The resulting belief in the conclusion is then a weighted sum of the degrees of belief in the premises:

$$bel'_c(C) = \sum_{i=1}^{n} bel_{\Rightarrow}(p_i \Rightarrow C) bel^i_p(p_i) \tag{4}$$

In front of these three expressions (equations (1), (2) and (4)) which result from the same logical form of the argument, a question arises. Which of them is the most appropriate for this argument type? The choices of the assignment of mass function in each equation do not model the same situation. Comparing the first and second formulas, on the one hand (1) suggests that each argument is based on one piece of evidence that could support the whole conclusion and is independent from the other ones. On the other hand, 2 supposes that all such arguments are provided at once by a single source. If we also compare equations (1) and (4), (1) represents the situation when the elementary arguments are independent, so that a mass is allocated each implication independently of the others. On the contrary, in (4) the belief mass assigned to in one implication affects those assigned to other ones, because, due to the use of a single mass function, the sum of all such masses must be one.

It is important to use each formula giving the belief in the conclusion in the appropriate situation, laying bare the underlying assumptions. For instance, the A-Arg presented in [4] is formally defined by $\wedge_{i=1}^{n}(p_i \Rightarrow C)$, and uses Yager's combination rule to calculate the overall confidence in the conclusion. However, Yager's rule was developed to deal with highly conflicting sources in place of Dempster rule. But the authors of [4] do not explain the presence of a conflict between pieces of evidence. Conflicts occur in cases when the intersection between focal sets is empty, which could be the case if masses were assigned to expressions supporting the negation of the implication (e.g., $p_i \wedge \neg C$), or in argument types involving equivalence. In the argument types discussed above, the focal sets resulting from handling argument of types 3, 4, 5 inspired by the selected articles do not generate such conflicts. In particular, the degree of disbelief in the conclusion is always 0 with these argument types.

4 Lessons learned

As shown in the previous sections, defining an "Argument Type" is a very delicate process. It depends on the assessor's understanding of the argument. Four important issues emerge from this paper:

Formal representation of the argument: The assessor should faithfully translate the informal definition of each argument into a formal one, by choosing the proper logical connectives relating premises to one another (conjunction,

disjunction) or between them and the conclusion (equivalence, implication). In order to do so, it is necessary to avoid vagueness and imprecision in the formulation of (informal) verbal definitions and to describe the characteristics of each argument type as accurately as possible. In this paper, two basic types have been laid bare : Those using conjunction of premises (Type 3) and those using a disjunction (Type 4 and 5). Indeed, it is important to know if for instance the truth in one premise is sufficient to ensure the conclusion, or if all premises are necessary to ensure the conclusion.

Using equivalence vs. implication connectives: Implication is the commonly used logical operator for representing arguments supporting a conclusion. However, we have seen that some verbal descriptions used in the literature on safety arguments can be formally translated into equivalences. The equivalence operator should be used carefully because it involves several situations that we may not intentionally want to encounter. Consider for example a conclusion (C) supported by one premise (P). Using the rule $(P \equiv C)$ implies that $(P \Rightarrow C)$ and that $(C \Rightarrow P)$. We saw in Section 3.4 that the second implication is not necessary true. In addition to this, equivalence is also expressed as $(\neg P \equiv \neg C)$, in particular, $\neg P \Rightarrow \neg C$); this is why using equivalence, disbelief in the conclusion may be different from zero.

Assigning mass functions: It should be clear that the formal definition of the argument types is not enough determine the degree of belief in the conclusion. For instance, when we changed the mass functions definition in the 4^{th} and 5^{th} logically equivalent types in Table 1 while using the same combination rule (i.e., Dempster rule), we obtained three distinct formulas (1, 2, 4). It is necessary to be sure that the choice of mass assignment reflects as well as possible the situation described in the arguments.

Choosing a combination rule : Changing the combination rule obviously affects the result of uncertainty propagation. So, it is important to choose the right one. However, as we saw in Section 3.4, applying Dempster combination rule is well adapted to computing the overall confidence, because no conflict is met during the combination step using arguments modelled by implication. In that case it is equivalent to Yager's rule. A possible use of this rule could be justified to cope with conflicts between the involved pieces of evidence, in place of Dempster rule. On the other hand, the disjunctive rule is too weak to be applied if one wants to jointly exploit the pieces of evidence in the safety case. A trade-off could be the rule in [11] which combines focal sets conjunctively when they are consistent and disjunctively when they conflict.

5 Conclusion

In this paper, we propose a comparative study between some confidence propagation methods in safety cases. We highlight four important elements to be considered in the development of a safety case. First, arguments should be expressed in formal logic. Second, we advocate the use of the implication connective, rather than equivalence, to describe the relationship between premises and conclusion.

Then, we propose a simplified framework to define mass functions attached to premises and arguments. Finally, we argue that Dempster rule of combination should be preferred when the focal sets issued from independent mass functions to be combined do not conflict.

In future works, we plan to experiment an approach that exploits this methodology in an application pervaded with high uncertainties, such as autonomous vehicles. Another issue is the improvement of methods proposed for translating expert opinions into usable numerical values, such as those proposed in [18] and applied in [4, 24]. In this regard, it would also be interesting to develop non-quantitative approaches using qualitative counterparts of belief functions as suggested in [9].

References

1. Ayoub, A., Chang, J., Sokolsky, O., Lee, I.: Assessing the overall sufficiency of safety arguments. In: 21st Safety-critical Systems Symposium (SSS'13), Bristol, United Kingdom (2013)
2. Bloomfield, R., Netkachova, K.: Building blocks for assurance cases. In: 2014 IEEE International Symposium on Software Reliability Engineering Workshops. pp. 186–191. IEEE (2014)
3. Chatalic, P., Dubois, D., Prade, H.: An approach to approximate reasoning based on Dempster rule of combination. Inter. J. of Expert Systems Research & Applications **1**, 67–85 (1987)
4. Cyra, L., Górski, J.: Support for argument structures review and assessment. Reliability Engineering & System Safety **96**(1), 26–37 (2011)
5. De La Vara, J.L., Génova, G., Álvarez-Rodríguez, J.M., Llorens, J.: An analysis of safety evidence management with the structured assurance case metamodel. Computer Standards & Interfaces **50**, 179–198 (2017)
6. Denney, E., Pai, G., Habli, I.: Towards measurement of confidence in safety cases. In: 2011 International Symposium on Empirical Software Engineering and Measurement. pp. 380–383. IEEE (2011)
7. Denoeux, T.: Conjunctive and disjunctive combination of belief functions induced by nondistinct bodies of evidence. Artif. Intell. **172**(2-3), 234–264 (2008)
8. Destercke, S., Dubois, D.: Idempotent conjunctive combination of belief functions: Extending the minimum rule of possibility theory. Inf. Sci. **181**(18), 3925–3945 (2011)
9. Dubois, D., Faux, F., Prade, H., Rico, A.: A possibilistic counterpart to shafer evidence theory. In: IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), New Orleans, LA, USA, June 23-26. pp. 1–6. IEEE (2019)
10. Dubois, D., Prade, H.: A set-theoretic view of belief functions. Logical operations and approximation by fuzzy sets. International Journal of General Systems **12**(3), 193–226 (1986)
11. Dubois, D., Prade, H.: Representation and combination of uncertainty with belief functions and possibility measures. Comput. Intell. **4**, 244–264 (1988)
12. EN50129: Railway applications - Safety related electronic systems for signaling (2003), CENELEC, European Committee for Electrotechnical Standardization
13. Graydon, P.J., Holloway, C.M.: An investigation of proposed techniques for quantifying confidence in assurance arguments. Safety science **92**, 53–65 (2017)

14. Guiochet, J., Do Hoang, Q.A., Kaaniche, M.: A model for safety case confidence assessment. In: International Conference on Computer Safety, Reliability, and Security (Safecomp). pp. 313–327. Springer (2014)
15. Hobbs, C., Lloyd, M.: The application of bayesian belief networks to assurance case preparation. In: Achieving Systems Safety, pp. 159–176. Springer (2012)
16. Inagaki, T.: Interdependence between safety-control policy and multiple-sensor schemes via Dempster-Shafer theory. IEEE Transactions on Reliability **40**(2), 182–188 (1991)
17. ISO 26262: Software considerations in airborne systems and equipment certification (2011), International Organization for Standardization (ISO)
18. Jøsang, A.: Subjective logic. Springer (2016)
19. Kelly, T.: Arguing Safety – A Systematic Approach to Safety Case Management. Ph.D. thesis, Department of Computer Science, University of York, UK (1998)
20. Kelly, T.P., McDermid, J.A.: Safety case construction and reuse using patterns. In: International Conference on Computer Safety, Reliability, and Security (Safecomp) 97, pp. 55–69. Springer (1997)
21. Sentz, K., Ferson, S., et al.: Combination of evidence in dempster-shafer theory. Tech. Rep. 0835, Sandia National Laboratories, Albuquerque, NM., USA (2002)
22. Smets, P.: Analyzing the combination of conflicting belief functions. Inf. Fusion **8**(4), 387–412 (2007)
23. Toulmin, S.E.: The Uses of Argument. Cambridge Univ. Press, Cambridge (1958)
24. Wang, R.: Confidence in safety argument-An assessment framework based on belief function theory. Ph.D. thesis, Institut National des Sciences Appliquées de Toulouse, France (2018)
25. Wang, R., Guiochet, J., Motet, G., Schön, W.: D-S Theory for Argument Confidence Assessment. In: 4th International Conference on Belief Functions (BELIEF 2016). pp. 190–200. Prague, Czech Republic (Sep 2016)
26. Wang, R., Guiochet, J., Motet, G., Schön, W.: Modelling Confidence in Railway Safety Case. Safety Science (110 part B), 286–299 (Dec 2018)
27. Wang, R., Guiochet, J., Motet, G., Schön, W.: Safety Case Confidence Propagation Based on Dempster-Shafer theory. International Journal of Approximate Reasoning **107**, 46–64 (Apr 2019)
28. Yager, R.R.: On the Dempster-Shafer framework and new combination rules. Information sciences **41**(2), 93–137 (1987)
29. Yuan, C., Wu, J., Liu, C., Yang, H.: A subjective logic-based approach for assessing confidence in assurance case. International Journal of Performability Engineering **13**(6), 807–822 (2017)