Uncertainty Elicitation and Propagation in GSN Models of Assurance Cases

Yassir Idmessaoud, Didier Dubois, Jérémie Guiochet

Uncertainty Elicitation and Propagation in GSN Models of Assurance Cases

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

[1] LAAS-CNRS, University of Toulouse, France
{yassir.id-messaoud,jeremie.guiochet}@laas.fr
[2] IRIT, University of Toulouse, France
dubois@irit.fr

Abstract. Goal structuring notation (GSN) is commonly proposed as a structuring tool for arguing about the high-level properties (e.g. safety) of a system. However, this approach does not include the representation of uncertainties that may affect arguments. Several works extend this framework using uncertainty propagation methods. The ones based on Dempster-Shafer Theory (DST) are of interest as DST can model incomplete information. However, few works relate this approach with a logical representation of relations between elements of GSN, which is actually required to justify the chosen uncertainty propagation schemes. In this paper, we improve previous proposals including a logical formalism added to GSN, and an elicitation procedure for obtaining uncertainty information from expert judgements. We briefly present an application to a case study to validate our uncertainty propagation model in GSN that takes into account both incomplete and conflicting information.

Keywords: Uncertainty propagation · Belief elicitation · Goal structuring notation · Dempster-Shafer application · Safety cases.

1 Introduction

Due to its expressiveness, the goal structuring notation (GSN) has became a defacto standard for graphical documentation of argument structures. It is notably used to argue about the safety of critical systems. However, even a well-designed GSN may include uncertainties that may question the final statement of the GSN. There is a lack of consensus about how to model these uncertainties in the argument structure. An interesting proposal [19] is to use Dempster-Shafer Theory (DST), since incomplete information can be explicitly modeled and calculated with. Several research works are investigating its use, but as presented in [7], the proposed uncertainty propagation schemes are often not clearly justified. This is mainly due to a lack of a clear definition of the logical relations between GSN elements. We investigate this issue in this paper, using DST and logical representations of arguments with new propagation models. We do not replace GSN informal notation, but build a formal model on top of it to propagate uncertainties. We also study how expert judgments can be elicited to feed our models.

The paper is structured as follows. Section 2 presents background and some related works. Sections 3 and 4 present the uncertainty propagation and elicitation methods respectively. Finally, Section 5 presents some experimental results gained by the proposed approach.

2 Background and related work

Goal structuring notation (GSN) is a graphical notation/language which represents argument structures (i.e., safety and assurance cases) in form of directed acyclic graphs (directed trees or arborescences). It breaks down a top claim, called "goal", into elementary sub-goals following a specific strategy and in accordance with a particular context. Each sub-goal is associated with pieces of evidence, called solutions, which support the conclusion. Figure 1 represents a typical hazard avoidance GSN pattern. To be considered as "*acceptably safe*" (G_1) all hazards (G_2 to G_n) of the system (X), listed in the context box (C_1), should be provably handled (Sn_1, Sn_2, ...) following the strategy (S_1). However, this symbol-based language does not specify the nature of the logical links between $G_1, G_2, \dots G_n$, nor does it capture the uncertainty that may exist in the argument structure. Previous works [5, 7] stated and discussed proposals that deal with the issue of uncertainty. An important part of these studies use probability theory to address it [4, 8]. For instance, some authors [15] transform GSN into a Bayesian network (BBN) and propagate probabilities accordingly. Due to the limited expressiveness of the probabilistic framework when information is lacking, such approaches can properly deal with uncertainties due to aleatory phenomena, but they poorly represent epistemic uncertainties due to incomplete information. In addition, these methods are also very greedy in terms of data, which requires much time to collect and process.

As a generalization of probability theory, Dempster-Shafer theory [16] (DST) offers tools to model and propagate both aleatory and epistemic uncertainty. A mass function, or basic belief assignment (BBA), is a probability distribution over the power set of the universe of possibilities (Ω), known as the *frame of discernment*. Formally, a mass function $m : 2^\Omega \rightarrow [0, 1]$ is such that $\sum_{E \subseteq \Omega} m(E) = 1$, and $m(\emptyset) = 0$. Any subset E of Ω such that $m(E) > 0$ is called a focal set of m. $m(E)$ quantifies the probability that we only know that the truth lies in E; in particular $m(\Omega)$ quantifies the amount of ignorance. A mass assignment induces a so-called belief function $Bel : 2^\Omega \rightarrow [0, 1]$, defined by: $Bel(A) = \sum_{E \subseteq A} m(E)$. It represents the sum of all the masses supporting a statement A. Belief in the negation $\neg A$ of the statement A is represented by: $Disb(A) = Bel(\neg A)$; the value $Uncer(A) = 1 - Bel(A) - Disb(A)$ quantifies the lack of information about A. In this paper, a *conjunctive rule of combination* is used for uncertainty propagation. This rule combines multiple pieces of evidence (represented by mass functions m_i, with $i = 1, 2, ..., n$) coming from independent sources of information: $m_\cap = m_1 \otimes m_2$ such that $m_\cap(A) = \sum_{E_1 \cap E_2 = A} m_1(E_1) \cdot m_2(E_2)$. In DST, an additional step eliminates conflicts that may exist by means of a normalization factor (dividing $m_\cap$ by

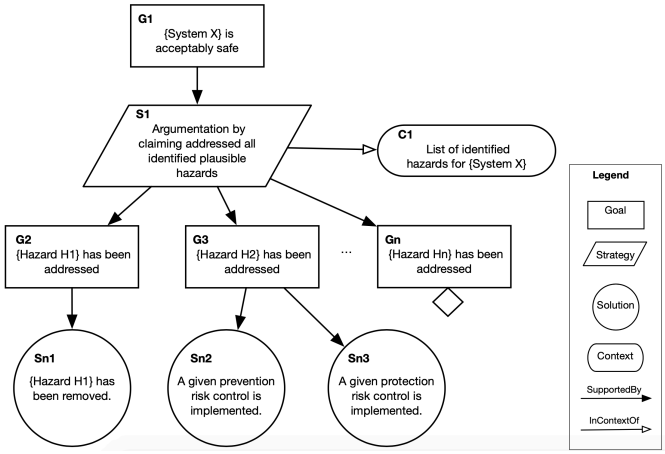

Fig. 1. GSN example adapted from the Hazard Avoidance Pattern [14]

$1 - m_{\cap}(\emptyset)$). This is Dempster rule of combination [16]. This step is omitted here to indicate the presence of possibly conflicting pieces of information.

Our approach builds on some previous works (mainly [2, 19]) that define a number of argument types and associate to each of them an uncertainty propagation formula in the setting of DST. However, in [2], no logical framework is provided, which prevents a formal justification of uncertainty propagation formulas. An implicit logical setting is offered in [19]. But it remains questionable since, for instance, rules that represent the relations between premises and the top-goal are modelled by equivalences. In our work we explicitly build propagation rules on a logical framework and we adopt a more flexible format using implications. A second issue is the elicitation process that collects information from experts and transforms it into belief and disbelief pairs in DST. For that, the method proposed in [2] and taken over in [19] is *ad hoc*. This transformation between expert information and (belief, disbelief) pairs is not *one to one* when the expert expresses no information. It yields some anomalous cases as discussed in [9]. Finally, no proposal was given in [2] to elicit belief on rules, while in [19] negative beliefs can be obtained, which is not acceptable. In this paper, we propose a new better-behaved elicitation approach based on the pignistic transform proposed in [17] that solves the two last issues.

3 From GSN to Dempster-Shafer Theory

As defined by [13, 14], Goal Structuring Notation (GSN) is a non-formal representation that does not formally specify how premises support a conclusion. In order to model such a relation, we use logical expressions. Then we shall attach degrees of uncertainty to these logical expressions and explain how to propagate these degrees of uncertainty in the GSN, in agreement with classical logic.

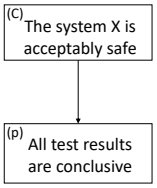

Fig. 2. A conclusion supported by one premise in GSN.

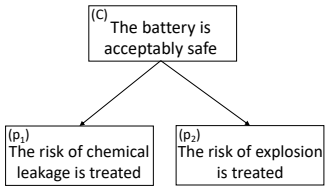

Fig. 3. A conclusion supported by two premise in GSN.

3.1 Logical modeling of GSN

Figure 2 represents a conclusion (C) supported by a single premise (p). It describes the situation in which the conclusion (C) is true if the premise (p) supporting it is also true. This statement can be expressed using a logical implication connective: $p \Rightarrow C$ standing for $\neg p \vee C$, using negation $\neg$ and disjunction $\vee$. It is obvious that such an expression can only assert the validity of the conclusion (in case p holds), i.e., whether C is provably true, not whether it is provably false. Note that even if C can only be true or false, we may fail to know it. So we work in a three-state universe (belief, disbelief and ignorance). To establish disbelief in C, we need to add an implication of the form $\neg p \Rightarrow \neg C$. It describes the situation where the conclusion (C) would be false, when the premise (p) is false. We call such logical expressions "rules". Those that induce belief in C are called *direct rules* and those that induce disbelief are called *reverse rules*.

With complex systems, it is more likely to find claims supported by more than one piece of evidence. In these cases, it is necessary to consider the relationship between the premises that support the same claim. On the other hand, logical implications remain the only connective that links the evidence domain to the conclusion. Through the different GSN patterns encountered in the literature, we can identify three types:

- **Conjunctive (C-Arg):** It describes the case when all premises are needed to support the conclusion. The direct rule is obtained by translating this definition into a logical expression: $(\wedge_i^n p_i) \Rightarrow C$. On the other hand, the reverse one is obtained by reversing the direct one: $\neg(\wedge_i^n p_i) \Rightarrow \neg C$, which is equivalent to $\wedge_i^n(\neg p_i \Rightarrow \neg C)$, a conjunction of simple rules.
- **Disjunctive (D-Arg):** It describes the case when one premise is enough to support the whole conclusion. The corresponding rules are: $\wedge_i^n(p_i \Rightarrow C)$ (direct), and $(\wedge_i^n \neg p_i) \Rightarrow \neg C$ (reverse).
- **Hybrid (H-Arg):** It describes the case where each premise supports the conclusion to some extent, but their conjunction does it to a larger extent. This rule type could be considered as a general type which includes the two previous ones. In fact, conjunctive and disjunctive types correspond to limit cases of the hybrid one.

Figure 3 represents an example of the conjunctive type. To assert that the battery is acceptably safe, all risks of chemical leakage and explosion should be

treated. It gives the expression: $(p_1 \wedge p_2) \Rightarrow C$. On the other hand, if one of the risks remains present we may assert that the battery is unsafe, which gives the expressions: $\neg p_1 \Rightarrow \neg C$ and $\neg p_2 \Rightarrow \neg C$.

All rules defined above will be used to build our uncertainty propagation model. Since the conjunctive and disjunctive types represent a special case of the hybrid one, we will only present the last one. However, it is simple to deduce their expressions from the general formula.

3.2 Uncertainty propagation model

In order to build our uncertainty propagation model, we define two kinds of parameters:

- Uncertainty on premises: It is modeled as a mass function on each premise of the argument: $< m_p^1, ..., m_p^n >$. m_p^i assigns a mass to the premise p_i, one on its negation ($\neg p_i$) and one on the tautology (Ω, representing ignorance) summing to 1.
- Uncertainty on rules: It is used to evaluate the impact of premises on a conclusion. We associate a simple support function [16] to each rule r of the argument type. Each simple support function consists in assigning a mass $m_r(r) = s$ to the rule and another one $m_r(\Omega) = 1 - s$ to the tautology, these weights summing to 1. The set of mass functions is formally defined as: $< m_\Rightarrow, m_\Rightarrow^i, m_\Leftarrow, m_\Leftarrow^i >$, where:
  $m_\Rightarrow$ and $m_\Leftarrow$ represent, respectively, direct and reverse conjunctive mass functions that assign support to rules $(\wedge_i^n p_i) \Rightarrow C$ and $(\wedge_i^n \neg p_i) \Rightarrow \neg C$, respectively.
  $m_\Rightarrow^i$, and $m_\Leftarrow^i$ respectively, assign support to elementary rules $p_i \Rightarrow C$ and $\neg p_i \Rightarrow \neg C$ occurring in the disjunctive type.

Using the conjunctive rule of combination presented in section 2, to merge the masses on the rules (conjunctive and disjunctive ones) with the masses on premises ($m_\cap = m_\Rightarrow \otimes m_\Leftarrow \otimes m_\Rightarrow^i \otimes m_\Leftarrow^i \otimes m_p^i$), we quantify the uncertainty on the conclusion C [1]. Since we work on a two-state frame of discernment for both premises $\Omega_p = \{p_i, \neg p_i\}$ and conclusion $\Omega_C = \{C, \neg C\}$, masses and (dis-)belief degrees on premises, rules and the conclusion are equal. For instance, $m_C(C) = Bel_C(C)$ and $m_C(\neg C) = Bel_C(\neg C) = Disb_C(C)$. We can prove the following results, by projecting $m_\cap$ on the universe $\Omega_C = \{C, \neg C\}$:

$$\begin{aligned} Bel_C(C) =& Bel_\Rightarrow([\wedge_{i=1}^n p_i] \Rightarrow C) \cdot \prod_{i=1}^{n} \{Bel_p^i(p_i) \cdot [1 - Bel_\Rightarrow^i(p_i \Rightarrow C)]\} \\ &+ \{1 - \prod_{i=1}^{n} [1 - Bel_p^i(p_i) \cdot Bel_\Rightarrow^i(p_i \Rightarrow C)]\} \end{aligned} \quad (1)$$

$$Disb_C(C) = Bel_{\Leftarrow}([\wedge_{i=1}^{n}\neg p_i] \Rightarrow \neg C) \cdot \prod_{i=1}^{n}\{Disb_p^i(p_i) \cdot [1 - Bel_{\Leftarrow}^i(\neg p_i \Rightarrow \neg C)]\}$$
$$+ \{1 - \prod_{i=1}^{n}[1 - Disb_p^i(p_i) \cdot Bel_{\Leftarrow}^i(\neg p_i \Rightarrow \neg C)]\} \quad (2)$$

Where:

- $Bel_C(C)$ (resp. $Disb_C(C)$): the degree of belief (resp. disbelief) in the conclusion C obtained by projection of $m_\cap$ on Ω_C.
- $Bel_p^i(p_i)$ (resp. $Disb_p^i(p_i)$): the degree of belief (resp. disbelief) in the i^{th} premise.
- $Bel_{\Rightarrow}([\wedge_{i=1}^{n}p_i] \Rightarrow C)$ (resp. $Bel_{\Leftarrow}^i(\neg p_i \Rightarrow \neg C)$): the degree of belief in the direct conjunctive rule (resp. i^{th} reverse rule).

We can notice that each formula (1) and (2) is the result of the summation of two terms. The first part expresses a generalized conjunction (the product), and the second part reflects a generalized disjunction (the probabilistic sum $1-(1-a)(1-b)$). To extract propagation formulas for the pure conjunctive type (C-Arg), it is enough to set to zero the masses on the direct rules ($Bel_{\Rightarrow}^i(p_i \Rightarrow C)$) and the mass on the conjunctive reverse rule ($Bel_{\Leftarrow}([\wedge_i^n\neg p_i] \Rightarrow \neg C)$). Similarly, to derive the pure disjunctive formulas (D-Arg), we set to zero the mass on the conjunctive direct rule ($Bel_{\Rightarrow}([\wedge_{i=1}^{n}p_i] \Rightarrow C)$) and the masses on the reverse rules ($Bel_{\Leftarrow}^i(\neg p_i \Rightarrow \neg C)$). We obtain:

$$\text{C-Arg}: \begin{cases} Bel_C(C) &= Bel_{\Rightarrow}([\wedge_{i=1}^{n}p_i] \Rightarrow C) \cdot \prod_{i=1}^{n} Bel_p(p_i) \\ Disb_C(C) &= 1 - \prod_{i=1}^{n}[1 - Disb_p^i(p_i) \cdot Bel_{\Leftarrow}^i(\neg p_i \Rightarrow \neg C)] \end{cases}$$

$$\text{D-Arg}: \begin{cases} Bel_C(C) &= 1 - \prod_{i=1}^{n}[1 - Bel_p^i(p_i) \cdot Bel_{\Rightarrow}^i(p_i \Rightarrow C)] \\ Disb_C(C) &= Bel_{\Leftarrow}([\wedge_{i=1}^{n}\neg p_i] \Rightarrow \neg C) \cdot \prod_{i=1}^{n} Disb_p^i(p_i) \end{cases}$$

Note that the belief (resp. disbelief) degree of the conclusion ($Bel_C(C)$) only depends on the belief (resp. disbelief) degree of premises ($Bel_p^i(p_i)$) and of the corresponding direct (reverse) rules ($Bel_{\Rightarrow}$ and $Bel_{\Rightarrow}^i$).

However, we observe in some cases that the sum of belief and disbelief of the conclusion, as calculated above, is greater than 1 which is not coherent. This is when the mass $m_\cap(\emptyset) > 0$. It is then counted in both sums defining the degrees of belief and disbelief. It indicates the presence of conflict between premises and rules. The coherence property $Bel_C(C) + Disb_C(C) \leq 1$ always hold if $m_\cap(\emptyset) = 0$. If it is not null, the conflict mass (3) should be subtracted from both belief and disbelief values, in order to get genuine contradiction-free degrees of belief and disbelief that respect the coherence property.

In [10], we provided a recursive equation to compute $m_\cap^n(\emptyset)$ for n premises when we know $m_\cap^{n-1}(\emptyset)$:

$$m_\cap^n(\emptyset) = Bel_C^{n-1}(C) \cdot m_n(\neg p_n \wedge \neg C) + Disb_C^{n-1}(C) \cdot m_n(p_n \wedge C) + m_\cap^{n-1}(\emptyset) \quad (3)$$

where:

- $Bel_C^{n-1}(C) = \{1 - \prod_{i=1}^{n-1}[1 - Bel_p^i(p_i) \cdot Bel_{\Rightarrow}^i(p_i \Rightarrow C)]\} - m_{\cap}^{n-1}(\emptyset)$
- $Disb_C^{n-1}(C) = \{1 - \prod_{i=1}^{n-1}[1 - Disb_p^i(p_i) \cdot Bel_{\Leftarrow}^i(\neg p_i \Rightarrow \neg C)]\} - m_{\cap}^{n-1}(\emptyset)$
- $m_i(p_i \wedge C) = Bel_p^i(p_i) \cdot Bel_{\Rightarrow}^i(p_i \Rightarrow C)$
- $m_i(\neg p_i \wedge \neg C) = Disb_p^i(p_i) \cdot Bel_{\Leftarrow}^i(\neg p_i \Rightarrow \neg C)$

Remark: D-Arg and C-Arg are conflict-free. Assuming that rule masses are maximal (= 1), for $n = 2$ we get: $Bel_C(C) = Bel_p^1(p_1) \cdot Bel_p^2(p_2)$ (for C-Arg) and $Bel_C(C) = Bel_p^1(p_1) + Bel_p^2(p_2) - Bel_p^1(p_1) \cdot Bel_p^2(p_2)$ (for D-Arg).

3.3 Belief and Disbelief elicitation

The model of uncertainty propagation presented above requires two types of inputs in order to compute belief and disbelief degrees of a conclusion: Belief/Disbelief on the rules and on the premises. These two information items will be directly collected from experts. To give their assessment about a premise or a rule, experts are asked to fill in an evaluation matrix, presented in Figure 4. Each point of this matrix corresponds to a strength of decision, denoted by

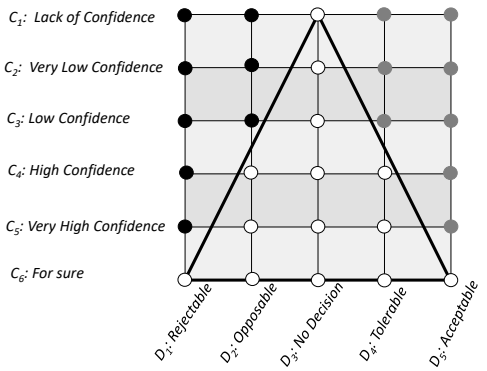

Fig. 4. Evaluation matrix

$Dec(A)$, and a degree of confidence in this decision, denoted by $Conf(A)$ attached to a proposition A. In a scale of 5 equidistant items, decision describes which side the expert leans towards: From the rejection ($Dec(A) = 0$) of a claim A, to its acceptance ($Dec(A) = 1$). It is formally the same as a degree of probability. On the other hand, confidence reflects the amount of information an expert possesses that can justify a decision. There are 6 equidistant levels of the confidence scale, from "Lack of confidence" $Conf(A) = 0$ to "For sure" $Conf(A) = 1$.

In Figure 5, we present four extreme expert assessments (see the black dot). The upper matrices represent the case of total confidence. The assessor rejects (resp. accepts) the claim in Figure 5.a (resp. 5.b). It corresponds to a maximal disbelief (resp. belief) degree. In contrast, the lower matrices represent resp. the

cases of total conflict (Figure 5.c) and ignorance (Figure 5.d). In both cases, the expert cannot make a clear decision either because he has as a lot of information both to support and reject the claim ($Conf(A) = 1$), or because he has no information ($Conf(A) = 0$). In contrast to other works [2, 18], we allow the assessor to use a midpoint value ($Dec(A) = 1/2$) to show full hesitancy.

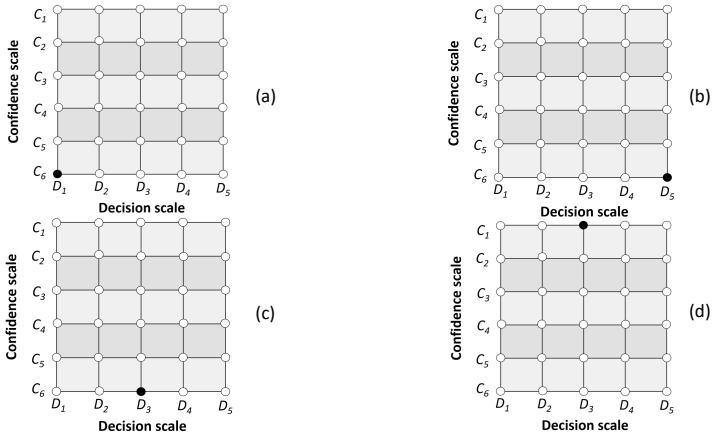

Fig. 5. Extreme assessments (black dot)

Uncertainty on premises: To be used in equations (1) and (2), the pair (decision, confidence) is translated into a triple (belief, disbelief and uncertainty). To do so, we use the formula proposed in [2], which defines confidence as the sum of belief and disbelief degrees (equation (4), left). On the other hand, we consider decision as the pignistic transform [17] that turns a mass into a probability (equation (4), right). So, we solve the following system for $Bel(p)$ and $Disb(p)$:

$$Conf(p) = Bel(p) + Disb(p); \quad Dec(p) = \frac{1 + Bel(p) - Disb(p)}{2} \tag{4}$$

However, as indicated in [9], the pignistic transform can generate negative belief and disbelief values when the pair $(Dec, Conf)$ given by an expert lies outside the triangle shown in Figure 4. Known as "Josang Triangle" [12], it represents a constraint that brackets decision $Dec(p)$ between two values:

$$\frac{1 - Conf(p)}{2} \leq Dec(p) \leq \frac{1 + Conf(p)}{2} \tag{5}$$

It guarantees that all clear-cut decisions (rejection or acceptance) are made only when the confidence level is maximal. To avoid negative belief and disbelief values, we must adjust the decision value to respect constraint (5). Therefore, when $Dec(p) < \frac{1-Conf(p)}{2}$ (rejection: black dots in Fig. 4), we set $Dec(p) = \frac{1-Conf(p)}{2}$. On the other hand, when $Dec(p) > \frac{1+Conf(p)}{2}$ (acceptance: grey dots in Fig. 4), we set $Dec(p) = \frac{1+Conf(p)}{2}$.

Example 1. Suppose we get the following assessments on two goals (p_1) and (p_2):

- p_1: Opposable with high confidence ($Dec(p_1) = 0.25, Conf(p_1) = 0.6$).
- p_2: Acceptable with very high confidence ($Dec(p_2) = 1, Conf(p_2) = 0.8$).

To calculate $Bel(p_i)$ and $Disb(p_i)$, we write them in terms of $Dec(p_i)$ and $Conf(p_i)$, from (4): $Bel(p) = \frac{Conf(p)-1}{2} + Dec(p)$, $Disb(p) = \frac{Conf(p)+1}{2} - Dec(p)$.

We can notice that the assessment for p_1 is inside the triangle in the matrix (Figure 4). Hence, there is no need to adjust the values:
$Bel(p_1) = \frac{0.6-1}{2} + 0.25 = 0.05$, $Disb(p_1) = \frac{0.6+1}{2} - 0.25 = 0.55$ and $Uncer(p_1) = 1 - Bel(p_1) - Disb(p_1) = 0.4$ for the amount of ignorance.

On the other hand, the assessment for p_2 is situated outside the triangle. In this case, we can be sure that the decision degree must be adjusted in accordance with the confidence value to get correct inputs. Before adjustment, we find a negative value of disbelief, which does not make sense: $Bel(p_2) = \frac{0.8-1}{2} + 1 = 0.9$ and $Disb(p_2) = \frac{0.8+1}{2} - 1 = -0.1$. Following the description above, we set $Dec(p_2) = \frac{1+Conf(p_2)}{2} = \frac{1+0.8}{2} = 0.9$. Then we find that $Bel(p_2) = 0.8$, $Disb(p_2) = 0$ and $Uncer(p_2) = 1 - Bel(p_2) - Disb(p_2) = 0.2$.

Uncertainty on rules: Assuming clear-cut knowledge about some (or all) premises ($Bel^i_p(p_i), Disb^i_p(p_i) \in \{0, 1\}$) and total ignorance about the others ($Uncer^i_p(p_i) = 1$), we notice that $Bel_C(C)$ and $Disb_C(C)$ in (1) and (2) are equal to rule masses. For instance, in the case of a conclusion C supported by two premises p_1 and p_2, assuming total acceptance of these two premises with maximal confidence, we get: $Bel_C(C) = Bel_{\Rightarrow}([p_1 \wedge p_2] \Rightarrow C)$. While assuming total rejection with maximal confidence of p_1, and total ignorance about p_2, we get: $Disb_C(C) = Bel_{\Leftarrow}(\neg p_1 \Rightarrow \neg C)$.

In order to collect masses on rules, under the assumption mentioned above (sure truth, sure falsity or ignorance on premises) we use the same approach as for eliciting uncertainty on premises. First, using the evaluation matrix (Figure 4), we take the expert opinions about the conclusion (which corresponds to the rules masses under those assumptions). Then, we change them to belief values using transformation formulas (4).

Moreover, we assume that a rule is either accepted or discarded, but not negated. In fact, for any rule $R : p \Rightarrow C$ we do not consider a positive disbelief because this would imply a belief in $\neg(p \Rightarrow C) = p \wedge \neg C$, i.e., $\neg R$ which is not a rule. So we only assign mass to a rule or to the tautology; the latter is the extent to which a rule is discarded. This constraint impacts the allowed pairs (Dec, Conf) for the expert. The latter is constrained to choose only a decision on the positive side (from "*No decision*" to "*acceptable*") for direct rules. On the contrary, (s)he can only choose a negative decision (from "*rejectable*" to "*No decision*") for the reverse rules. Formulas in (4) are used to derive the degrees of belief on rules.

Example 2. Consider the case of Figure 2:

- Direct rule ($R_1 : p \Rightarrow C$): Assuming $Dec(p) = 1$, expert assigns "Tolerable with high confidence" to C: $Dec(C) = 0.75,\ Conf(C) = 0.6$
- Reverse rule ($R_2 : \neg p \Rightarrow \neg C$): Assuming $Dec(p) = 0$, expert assigns "Opposable with very high confidence" to C: $Dec(C) = 0.25,\ Conf(C) = 0.8$

We can notice in this example that both cases respect the Josang constraint (5). Hence, there is no need to adjust the decision value. Using (4) for the direct rule R_1: $Bel_{\Rightarrow}(R_1) = Bel_C(C) = \frac{(0.6)-1}{2} + (0.75) = 0.55$ and we set $Bel_{\Rightarrow}(\neg R_1) = 0$. In the same way, for the reverse rule R_2: $Bel_{\Leftarrow}(R_2) = Disb_C(C) = \frac{(0.8)+1}{2} - (0.25) = 0.65$ and we set $Bel_{\Leftarrow}(\neg R_2) = 0$.

4 Uncertainty assessment procedure

In this section, we present our approach to uncertainty propagation from premises to the top goal of a GSN. As illustrated on Figure 6, this procedure is structured in two phases.

The first one, called *modeling phase*, collects expert opinions on rules, expressed with qualitative scores $(Dec, Conf)$, and translates them into numerical mass assignments to rules. It will be conducted by asking $(2n + 2)$ questions to the assessor using the evaluation matrices, n being the number of premises. The first $(2n)$ questions concern masses on elementary rules (direct and reverse). For instance, to get, respectively, the values of $Bel^i_{\Leftarrow}(\neg p_i \Rightarrow \neg C)$ and $Bel^i_{\Rightarrow}(p_i \Rightarrow C)$ the expert will be asked the following questions (in case $n = 2$):

1. Supposing no knowledge about premise p_1 (resp. p_2) : $(Dec = 0.5, Conf = 0)$ and minimal Dec value (rejectable for sure) in premise p_2 (resp. p_1): $(Dec = 0, Conf = 1)$, what is your Decision/Confidence in the conclusion?
2. Supposing no knowledge about premise p_1 (resp. p_2): $(Dec = 0.5, Conf = 0)$ and a maximal Dec value (acceptable for sure) concerning premise p_2(resp. p_1) : $(Dec = 1, Conf = 1)$, what is your Decision/Confidence in the conclusion?

The additional two questions concern the conjunctive rules (resp. reverse and direct):

3. Supposing minimal Dec value (rejectable for sure) concerning both premises p_1, p_2 : $(Dec = 0, Conf = 1)$, what is your Decision/Confidence in the conclusion?
4. Supposing maximal Dec value (acceptable for sure) concerning both premises p_1, p_2 : $(Dec = 1, Conf = 1)$, what is your Decision/Confidence in the conclusion?

We assume that once these masses on rules are evaluated, they can be used for the considered system using the second phase explained below.

The second phase, called *application phase*, concerns the collection of expert data on premises. One question per premise is then formulated to the experts: considering the knowledge on the pieces of evidence (also called *solutions* in GSN), what is your "Decision" and "Confidence" concerning premise p_i?

Grouped in a questionnaire, these $(3n+2)$ questions will be asked in form of matrices to be filled in by the assessor (for rules, some matrices may be pre-filled, see Figure 6). Then, these values (on rules and premises) are used to calculate the belief/disbelief in the conclusion (eqs. (1) and (2)). Finally, we may transform the resulting triple (Belief, Disbelief, Uncertainty) concerning the conclusion, to a pair (Decision, Confidence) using formulas (4) and approximate them by qualitative values.

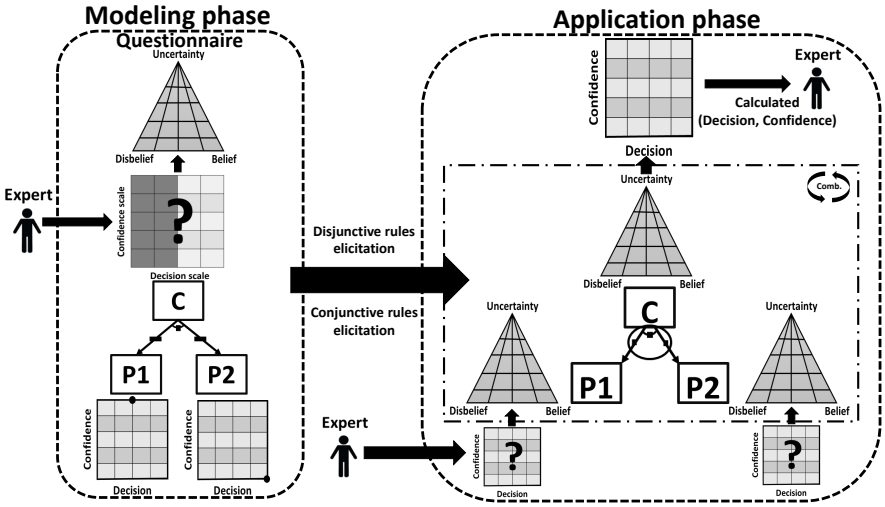

Fig. 6. Schema of the assessment framework for safety argument

5 Case study

In this section, we use a portion of GSN proposed in [3] to test and validate our uncertainty propagation approach. That study proposed a hybrid architecture of a collision avoidance system for drones, Urban Air Mobility and Air Taxis with horizontal automatic resolution. It is named ACAS-X (Next-Generation Airborne Collision Avoidance System). It replaces a set of lookup tables (LUTs) (that provide anti-collision maneuvering guidance according to the speed of the two aircrafts, their relative positions, and the time until the loss of vertical separation occurs) by a neural network (NN) of much smaller size. In addition to the NN-based controller, this architecture includes a safety net which contains a portion of LUTs (already established as safe) for unsafe areas (where the NN may give results different from those of the LUTs), and a check module which controls the switch between these two sub-systems (NN and LUTs). The GSN section (figure 7) in which we are interested, argues that "G_1: *Real world situations where MLM*[3] *is not robust are identified and mitigated*". To demonstrate this statement, the top goal (G_1) is broken down into two sub-goals (G_2) and (G_3). (G_2) ensures that the property was correctly defined to identify all unsafe

[3] Machine learning Model.

situations (G_4) and formally checked (G_6) in each of the areas (called p-boxes) into which the input space was correctly decomposed (G_5). (G_3) ensures that unsafe situations were properly mitigated via an appropriate architecture (G_7).

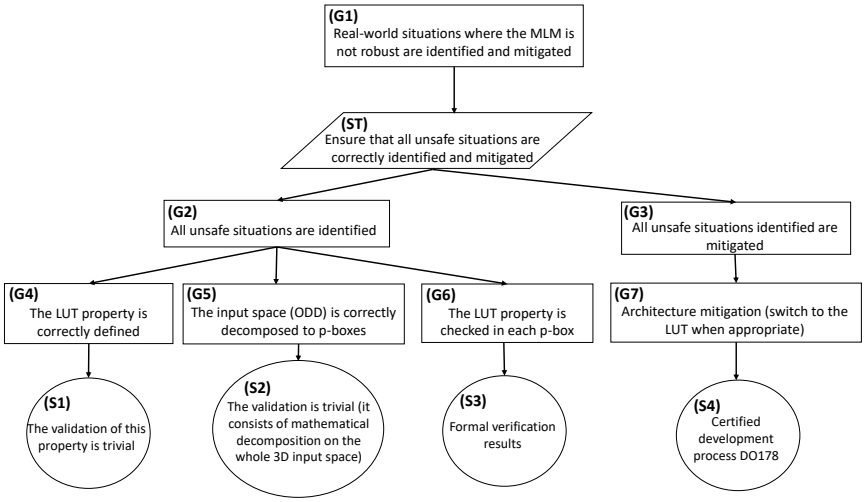

Fig. 7. Assurance Case - ML subsystem robustness [3]

Table 1 groups the degrees of belief on the rules involved in this case. Following the assessment procedure above, these values are the result of a questionnaire[4] answered by a safety expert about this system. We can notice that all direct conjunctive rules receive maximal weights and the elementary rule weights for (G_1) and (G_2) are null. Thus, we deduce that this GSN represent a conjunctive type where all sub-goals are needed to support (G_1). As seen in [10], C-Arg tends to propagate the premises that support the conclusion with the least weight, increasing along with it the uncertainty level. Thus, we can explain why we go from acceptable premises with very high confidence (G_6, G_7), high confidence (G_5) and for sure (G_4) to a tolerable top goal (G_1) with low confidence ($Dec = 0.692, Conf = 0.384$). Graphs in figures 8 and 9 present, respectively, the sensitivity of decision and confidence degrees of the conclusion (G_1) to the sub-goals (G_4), (G_5), (G_6) and (G_7). To determine the latter, we vary the value of a premise from its minimal to its maximal value, while we keep the values of the other premises to their base values. We can notice that all values, are indeed included in the interval [0,1]. We can also notice that the pair (decision, confidence) on the goal (G_1) varies from *"Rejectable for sure"* ($Dec = 0, Conf = 1$) to *"Tolerable with high confidence"* ($Dec = 0.82, Conf = 0.64$). The sub-goal (G_4) has the lowest influence on decision and the highest influence on confidence; the opposite applies for sub-goal (G_5).

[4] The questionnaire is available in [11].

Table 1. Elicited belief degrees on rules

Goal (G_i)	Belief degree on rules
G_1 $(i = 2, n = 3)$	$Bel_{\Rightarrow}([\wedge_i^n G_i] \Rightarrow G_1) = 1$ $Bel_{\Leftarrow}([\wedge_i^n \neg G_i] \Rightarrow \neg G_1) = 1$ $Bel_{\Rightarrow}(G_i \Rightarrow G_1) = 0$ $Bel_{\Leftarrow}(\neg G_i \Rightarrow \neg G_1) = 1$
G_2 $(i = 4, n = 6)$	$Bel_{\Rightarrow}([\wedge_i^n G_i] \Rightarrow G_2) = 1$ $Bel_{\Leftarrow}([\wedge_i^n \neg G_i] \Rightarrow \neg G_2) = 1$ $Bel_{\Rightarrow}(G_i \Rightarrow G_2) = 0$ $Bel_{\Leftarrow}(\neg G_i \Rightarrow \neg G_2) = 1$
G_3	$Bel_{\Rightarrow}(G_7 \Rightarrow G_3) = 1$ $Bel_{\Leftarrow}(\neg G_7 \Rightarrow \neg G_3) = 1$

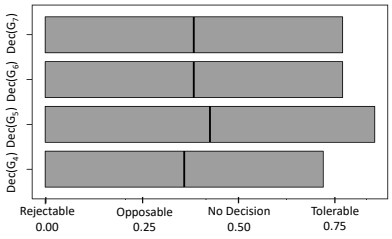

Fig. 8. Decision sensitivity on the top goal G_1

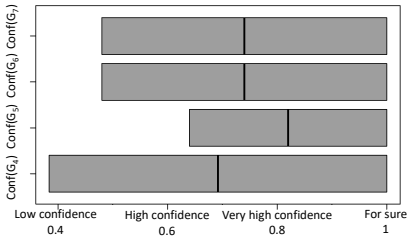

Fig. 9. Confidence sensitivity on the top goal G_1

6 Conclusion

In this paper, we propose an extensive approach to the elicitation and propagation of uncertainty in a logical GSN model and report on a preliminary case study for testing our approach. However, some issues still need to be addressed. First of all, our propagation model does not consider all GSN components (such as Justification, Assumption, etc.). In addition, our elicitation model seems to encourage experts to give extreme values of (decision, confidence) so that we often end up with a conjunctive or disjunctive type. But these two types are not the only types that exist in literature. Finally, the transformation of expert opinion from quantitative to qualitative values is also a source of uncertainty. In a future work, we plan to develop a purely qualitative approach to information fusion based on [6], and compare it to the quantitative one.

Acknowledgement

A special thanks to the authors of [3], especially to Christophe GABREAU for answering the questionnaire concerning the assessment of the GSN presented in our case study.

References

1. 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)
2. Cyra, L., Górski, J.: Support for argument structures review and assessment. Reliability Engineering & System Safety **96**(1), 26–37 (2011)
3. Damour, M., Grancey, F.D., Gabreau, C., Gauffriau, A., Ginestet, J.B., Hervieu, A., Huraux, T., Pagetti, C., Ponsolle, L., Clavière, A.: Towards certification of a reduced footprint acas-xu system: A hybrid ml-based solution. In: International Conference on Computer Safety, Reliability, and Security. pp. 34–48. Springer (2021)
4. 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)
5. Duan, L., Rayadurgam, S., Heimdahl, M.P., Ayoub, A., Sokolsky, O., Lee, I.: Reasoning about confidence and uncertainty in assurance cases: A survey. Software Engineering in Health Care pp. 64–80 (2014)
6. 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)
7. Graydon, P.J., Holloway, C.M.: An investigation of proposed techniques for quantifying confidence in assurance arguments. Safety science **92**, 53–65 (2017)
8. Guiochet, J., Do Hoang, Q.A., Kaaniche, M.: A model for safety case confidence assessment. In: International Conference on Computer Safety, Reliability, and Security. pp. 313–327. Springer (2014)
9. Idmessaoud, Y., Dubois, D., Guiochet, J.: Belief functions for safety arguments confidence estimation: A comparative study. In: International Conference on Scalable Uncertainty Management. pp. 141–155. Springer (2020)
10. Idmessaoud, Y., Dubois, D., Guiochet, J.: Quantifying confidence of safety cases with belief functions. In: International Conference on Belief Functions. pp. 269–278. Springer (2021)
11. Idmessaoud, Y., Guiochet, J., Dubois, D.: Questionnaire for estimating uncertainties in assurance cases (Apr 2022), https://hal.laas.fr/hal-03649068
12. Jøsang, A.: Subjective logic. Springer (2016)
13. Kelly, T.: Arguing Safety – A Systematic Approach to Safety Case Management. Ph.D. thesis, Department of Computer Science, University of York, UK (1998)
14. 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)
15. Nešić, D., Nyberg, M., Gallina, B.: A probabilistic model of belief in safety cases. Safety science **138**, 105187 (2021)
16. Shafer, G.: A mathematical theory of evidence. Princeton university press (1976)
17. Smets, P.: Decision making in the tbm: the necessity of the pignistic transformation. International Journal of Approximate Reasoning **38**, 133–147 (2005)
18. Wang, R., Guiochet, J., Motet, G.: Confidence Assessment Framework for Safety Arguments. In: International Conference on Computer Safety, Reliability, and Security (SAFECOMP), Trento, Italy. p. 14p. (Sep 2017)
19. 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)