

Uncertainty assessment in safety argument structure : an approach based on Dempster-Shafer theory

Yassir Idmessaoud

▶ **To cite this version:**

Yassir Idmessaoud. Uncertainty assessment in safety argument structure : an approach based on Dempster-Shafer theory. Cryptography and Security [cs.CR]. Université Paul Sabatier - Toulouse III, 2022. English. NNT : 2022TOU30263 . tel-04048943v2

HAL Id: tel-04048943

https://theses.hal.science/tel-04048943v2

Submitted on 28 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ FÉDÉRALE TOULOUSE MIDI-PYRÉNÉES

Délivré par :
l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)

Présentée et soutenue le *15/12/2022* **par :**

Yassir ID MESSAOUD

Uncertainty Assessment in Safety Argument Structure – An approach based on Dempster-Shafer Theory

---

JURY

Frédéric PICHON	Professeur des universités	Rapporteur
Walter SCHON	Professeur des universités	Rapporteur
Sylvie LE HEGARAT	Professeure des universités	Examinatrice
Jean-Loup FARGES	Maître de recherche ONERA	Examinateur
Jérémie GUIOCHET	Professeur des universités	Directeur de thèse
Didier DUBOIS	Directeur de recherche CNRS	Directeur de thèse

École doctorale et spécialité :
EDSYS : Informatique 4200018
Unité de Recherche :
Laboratoire d'Analyse et d'Architecture des Systèmes
Directeur(s) de Thèse :
Jérémie GUIOCHET et *Didier DUBOIS*
Rapporteurs :
Frédéric PICHON et *Walter SCHON*

Acknowledgments

The work presented in this thesis has been realised in the Laboratory for Analysis and Architecture of Systems (Laas), a laboratory depending from the French National Center for Scientific Research (CNRS). I would like to thank, first of all, our team head Madam Hélène WAESELYNCK and all the members of the team TSF (Dependable Computing and Fault Tolerance) for having welcomed me to carry out my work in this laboratory.

I would like furthermore express my sincere gratitude to my Ph.D supervisors, Dr. Didier DUBOIS from IRIT (Institute of Research in Computer Science of Toulouse) and Pr. Jérémie GUIOCHET from Laas-CNRS, my supervisors of the Ph.D. thesis, for their tremendous support during these three years. I could not have accomplished this ambitious mission without their advice, encouragement, patience and the time devoted in regular exchanges. I feel honored to have had the opportunity to work with both of them. I would also like to thank the rest of my thesis defense committee:

- Pr. Frédéric PICHON, Artois University, Reviewer
- Pr. Walter SCHON, UTC, Reviewer
- Pr. Sylvie LE HEGARAT, Paris saclay university, Examiner
- Dr. Jean-Loup FARGES, ONERA, Examiner

I really appreciate their acceptance to review my work, as well as their insightful feedback and interesting questions during my defense to improve the manuscript and expand my research from various perspectives.

I am also very grateful to Dr. Claire PEGATTI from ONERA and Christophe GABREAU from AIRBUS for their assistance on the case study used for the validation of our approach.

I would like to give my special thanks to the colleagues in TSF, especially the doctoral students, with whom I went through these unforgeable three years of the Ph.D. study. Specific appreciations are sent to my comrades: Yuxiao, Mohamed, Florent, Cyrius, Raul and Luca.

Contents

List of figures **vi**

List of Tables **viii**

Introduction **1**

1 General background and State of the art **5**
1.1 Introduction 5
1.2 Argument structures 6
1.2.1 A definition of argument structures 6
1.2.2 Forms of argument structures 7
1.3 An introduction to uncertainty theories 10
1.3.1 The concept of uncertainty 10
1.3.2 Uncertainty modeling 11
1.4 DST applied to confidence propagation 17
1.4.1 DS theory and logic 17
1.4.2 Graphical representation of belief propagation 18
1.5 Quantitative confidence propagation approaches in argument structure 19
1.5.1 Uncertainty theories in arguments 19
1.5.2 Argument type and logic 19
1.5.3 Expert elicitation 23
1.6 Conclusion 24

2 Uncertainty propagation model in argument structure **27**
2.1 Introduction 27
2.2 Confidence measure in an argument structure 28
2.2.1 Formal definition of uncertainty in premises 29
2.2.2 Formal definition of uncertainty in the support relation . . . 29
2.3 Confidence propagation models 30
2.3.1 Simple argument type (S-Arg) 30
2.3.2 Conjunctive argument type (C-Arg) 32
2.3.3 Disjunctive argument type (D-Arg) 35
2.3.4 Hybrid argument type (H-Arg) 37
2.3.5 Conflict mass for (n) premises 43
2.4 Argument types in literature - A comparison 46
2.5 Models sensitivity analysis 48
2.6 Conclusion 52

3 Expert opinion elicitation **55**
3.1 Introduction . . . 55
3.2 Elicitation model . . . 56
3.3 Belief and disbelief elicitation of premises . . . 57
3.4 Belief elicitation for rules . . . 62
3.5 Argument type identification . . . 64
3.6 Pignistic transform Vs Shenoy transform . . . 65
3.7 Conclusion . . . 67

4 Applying confidence assessment of an argument in GSN - Case study **69**
4.1 Introduction . . . 69
4.2 Confidence assessment procedure . . . 70
4.2.1 Modelling phase: Rules elicitation . . . 70
4.2.2 Application phase: Premises elicitation and confidence propagation . . . 73
4.3 Case study . . . 75
4.4 Discussion . . . 81
4.5 Conclusion . . . 82

5 An opening towards a pure qualitative approach **83**
5.1 Introduction . . . 83
5.2 Introduction to qualitative capacities . . . 84
5.3 Qualitative confidence propagation models . . . 86
5.3.1 Simple argument (S-Arg) . . . 86
5.3.2 Conjunctive argument (C-Arg) . . . 87
5.3.3 Disjunctive argument (D-Arg) . . . 88
5.3.4 Hybrid argument (H-Arg) . . . 89
5.3.5 Conflict mass . . . 91
5.4 Qualitative elicitation models . . . 92
5.5 Qualitative assessment application to GSN arguments . . . 93
5.6 Quantitative Vs Qualitative assessment procedure . . . 95
5.7 Discussion . . . 96
5.8 Conclusion . . . 97

Conclusion **99**

A Questionnaire for expert elicitation **105**

Bibliography **117**

List of figures

1.1 An example of a textual argument [Kelly 2004] 7
1.2 GSN main components 8
1.3 GSN example adapted from the Hazard Avoidance Pattern [Kelly 1997] 8
1.4 An example of a claim-argument-evidence (CAE) notation 9
1.5 Relation between uncertainty theories 14
1.6 An example of a valuation network 19
1.7 Argument type proposed by Annaheed et al. [Ayoub 2013] 20

2.1 Sources of uncertainty in GSN - A simple argument type (S-Arg) example (Note that the arrow direction is not intuitive regarding the implication $p \Rightarrow C$) 29
2.2 A conjunctive argument type (C-Arg) example 33
2.3 A disjunctive argument type (D-Arg) example 35
2.4 A hybrid argument type (H-Arg) example 37
2.5 Argument types 48
2.6 Belief sensitivity for a C-Arg 50
2.7 Belief sensitivity for a D-Arg 50
2.8 Belief sensitivity for a H-Arg 50
2.9 Disbelief sensitivity for a C-Arg 50
2.10 Disbelief sensitivity for a D-Arg 50
2.11 Disbelief sensitivity for a H-Arg 50
2.12 Uncertainty sensitivity for a C-Arg 51
2.13 Uncertainty sensitivity for a D-Arg 51
2.14 Uncertainty sensitivity for a H-Arg 51
2.15 Conflict degree sensitivity 51
2.16 Masses on rules and premises 53

3.1 The positioning of the uncertainty interval on the pair $(Dec, Conf)$. 57
3.2 Evaluation matrix 57
3.3 Extreme assessments (black dot) 58
3.4 Evaluation matrix including Josang constraint 59
3.5 Belief and disbelief elicitation of premises 61
3.6 Belief elicitation of rules 63
3.7 Responses from experts vs. Calculated results based on Complementary and Redundant argument types [Wang 2018b] 64
3.8 Pignistic (plain line) and Shenoy (dashed line) transforms constraint 66

4.1 Uncertainty assessment procedure 70
4.2 Modelling phase 71
4.3 Disjunctive reverse rule elicitation 72

4.4 Disjunctive direct rule elicitation 72
4.5 Conjunctive reverse rule elicitation 72
4.6 Conjunctive direct rule elicitation 72
4.7 Application phase 74
4.8 Premises elicitation 74
4.9 Architecture of the neural network based ACAS Xu [Damour 2021] . 76
4.10 ACAS Xu geometry [Katz 2017] 76
4.11 Assurance Case - ML subsystem robustness [Damour 2021] 77
4.12 Argument fragments according to their type 80
4.13 Decision sensitivity on the top goal G_1 80
4.14 Confidence sensitivity on the top goal G_1 80
4.15 Impact of changing the scale - example 81

5.1 Evolution of certainty and information in pairs (belief, disbelief) . . 85
5.2 Schema of the assessment framework for safety argument 94
5.3 Evaluation matrix 95
5.4 GSN artificial example 95

A.1 Evaluation Matrix 105
A.2 Example 1 106
A.3 Example 2 107
A.4 Assurance Case - ML subsystem robustness [Damour 2021] 107

List of Tables

1.1 DS combination rule example 16
1.2 Propagation formulas from [Cyra 2011] 21
1.3 Propagation formulas from [Wang 2019] 22

2.1 Combination of direct ($m_{\Rightarrow}$) and reverse ($m_{\Leftarrow}$) rules for S-Arg . . . 31
2.2 Combination of the mass on premise (m_p) with its rules (m_r) for S-Arg 31
2.3 Combination of the masses on premises 32
2.4 Combination of elementary reverse rules for the C-Arg 33
2.5 Combination of direct and reverse rules for the C-Arg 33
2.6 Combination of masses of rules and premises for the C-Arg (part 1) 33
2.7 Combination of masses of rules and premises for the C-Arg (part 2) 34
2.8 Combination of elementary direct rules for the D-Arg 35
2.9 Combination of direct and reverse rules for the D-Arg 36
2.10 Combination of masses of rules and premises for the D-Arg (part 1) 36
2.11 Combination of masses of rules and premises for the D-Arg (part 2) 36
2.12 Combination of direct and reverse conjunctive rules for H-Arg 38
2.13 Combination of direct and reverse elementary rules for the H-Arg . . 38
2.14 Combination of premise (p_i) with its elementary rules for H-Arg . . 38
2.15 Combination of premise (p_i) with its elementary rules of two premises (1 and 2) for H-Arg (part 1) 38
2.16 Combination of premise (p_i) with its elementary rules of two premises (1 and 2) for H-Arg (part 2) 39
2.17 Combination of premise (p_i) with its rules for H-Arg 40
2.18 Values of the 3-tuple (Bel, $Disb$, $Uncer$) resp. for C-Arg, D-Arg and H-Arg 48
2.19 Rules and premises basic values for the sensitivity analysis 49

3.1 Values from (Dec, $Conf$) to (Bel,$Disb$) couples on premises (see figure 3.2 for symbol meaning) 62

4.1 Decision items and their appropriate quantitative counterpart 75
4.2 Confidence items and their appropriate quantitative counterpart (see figure 3.2 for symbols meaning) 75
4.3 Elicited belief degrees on rules 78
4.4 Basic assessment values of premises ($Bel_p(G_i), Disb_p(G_i), Uncer_p(G_i)$) 79
4.5 Belief and disbelief degrees of (sub-)goals calculation 79
4.6 Impact of changing the decision scale on confidence assessment . . . 81

5.1 Combination of the focal sets of the premise (ρ_p) with its rules (ρ_r) for S-Arg 86

5.2 Combination of masses of rules and premises for the C-Arg (part 1) 87
5.3 Combination of masses of rules and premises for the C-Arg (part 2) 87
5.4 Combination of masses of rules and premises for the D-Arg (part 1) 88
5.5 Combination of masses of rules and premises for the D-Arg (part 2) 89
5.6 Combination of premise (p_i) with its rules for H-Arg 90
5.7 Values from (Dec, $Conf$) to (Bel,$Disb$) pairs on premises 93
5.8 Values of the belief degrees on rules 95
5.9 Pairs (decision, confidence) according to both qualitative (Qual.) and quantitative (Quant.) methods for the example (see Fig. 5.3 for the meaning of symbols) 96

Introduction

The trustworthiness of cyber-physical systems has become an important issue in their development process involving many professionals in the industrial and research sectors. Given its importance, many projects and initiatives, with considerable budgets, are being launched to meet these demands. The development of complex, autonomous and intelligent systems requires a significant effort to prove their dependability to meet governments standards or convince skeptical consumers. Indeed these systems, in particular the safety critical ones, usually operate in uncertain environments and include new technologies not yet fully understood (e.g., those involving machine learning). Hence developing assessment tools for these new types of systems is of paramount importance.

Many dependability techniques are useful to meet the requirements standards, including failure probability calculations. These pre-established safety techniques are well adapted to known contexts, but are insufficient to validate these new systems (e.g., autonomous vehicles), partly because they do not explicitly represent uncertainty related to lack of information when standards are not applicable, or when probability measures cannot be used. A new approach is under development known as: *structured arguments*. They are proposed to justify the satisfaction of a high-level property of a system. They are also called *assurance cases*. When the property to defend is safety, they are usually called *safety cases*. These cases could be based on structured languages even with graphical notations (e.g., *GSN*, for *Goal Structuring Notation*). Basic elements such as goal or evidence are the building blocks of a structure (an argument) in order to support a "top goal", or a conclusion (e.g., "the system is acceptably safe"). However, assurance cases only provide a qualitative analysis, and no estimation of the uncertainties that may exist on the evidence could be specified. This is a serious drawback, particularly to manage complexity, or to convince third parties (like certification bodies). In order to provide solutions to quantify uncertainty/confidence, many works augment these arguments (generally represented with graphical notations) with probabilistic models for confidence propagation. These models 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 continuation of previous studies, reviewed in this manuscript, we develop here a method for uncertainty assessment in argument structure based on *Dempster-Shafer Theory* (DST). This theory offers resources to model and propagate both aleatory and epistemic uncertainty. DST could be viewed as a generalisation of probability theory, where metrics such as belief, disbelief and uncertainty on a statement are explicitly quantified. This theory has been introduced in the 1960's as a mechanism for statistical reasoning under epistemic (knowledge) uncertainty by Arthur Dempster and developed later in the 1970's by Glenn Shafer with a more

subjectivist spirit. Since then, an important work has been carried out to provide mathematical tools to manipulate these DST metrics. Thus, in order to apply them in assurance cases, three main challenges can be identified:

- Uncertainty/Confidence quantification

  Identifying potential sources of uncertainty in an argument and quantify them is a first important challenge. Two sources of uncertainty can be identified in an assurance case. The first is related to the "evidence" collected for the case, impacting their trustworthiness. The second source is related to the argument structure itself. For instance, one may question the appropriateness of the "evidence" to the "top goal".

  To address this challenge, we use propositional logic to formally represent the relation between evidence and goal by means of "*if-then rules*". Then we used *belief functions* (presented in chapter 1) to quantify the uncertainty in the evidence and in the rules.

- Propagation formulas

  Once uncertainties in an argument are quantified (for all pieces of evidences for instance), it is important to propagate the uncertainties to the "top goal" in order to deduce the confidence in it.

  A combination rule that will be applied to merge the uncertainties must be defined. Indeed, in an argument the pieces of evidence may be redundant, or complementary, which will be completely different in terms of uncertainty propagation. This choice is based on several criteria. For example, the way of managing a conflict situation when two contradictory pieces of information needs to be merged. One may adopt an optimistic approach by assuming that the truth lies in one of the two sources. In this case, we can consider a disjunctive rule of combination. On the contrary, a more risky approach might be adopted that rejects or ignores contradictory information. In this case, we may consider conjunctive combination rules.

  After quantifying the sources of uncertainty in the assurance case, we identified an appropriate combination rule to build our confidence propagation model.

- Elicitation procedure

  Once the confidence propagation model is built, we need to provide the inputs to run it. Such inputs are often based on subjective judgements provided by experts and expressed in natural language (e.g., high, moderate or low confidence, etc). Hence the need to develop models to transform these judgments into quantitative measures of confidence without altering them. A second issue concerns the procedure for collecting these data. Questionnaires are often used to collect testimonials. However, since we are handling subjective information, it is crucial to be as explicit as possible when we survey experts so that the information is consistent with our expectations.

To deal with this challenge, we propose an expert judgment collection method using a questionnaire. We also propose formulas for transforming these collected judgments into quantitative measures of belief based on DST.

A fourth objective has been set during this thesis which aims at exploring theories that propose tools for the modeling and propagation of qualitative measures of confidence.

- Purely qualitative confidence assessment approach

  In so-called quantitative approaches to confidence propagation in assurance cases we often have to transform qualitative testimonies into numerical measures, in the $[0, 1]$ scale, to be used by quantitative propagation models. This transformation is somewhat arbitrary, especially when we approximate quantitative confidence in the "top goal" (after the propagation) with a qualitative scale.

  To address this issue, we outline an alternative to DST, known as *Qualitative Capacity Theory* (QCT), to qualitatively assess and propagate confidence in assurance cases.

We address in this thesis these four challenges with theoretical contributions applied to artificial examples but also to a case study in the aeronautics field. This thesis manuscript will be organised as follows:

In chapter 1, we present the theoretical background and review the bibliography of the works related to the thesis. It is structured in four parts. First of all, we introduce the concept of argument structure and present its different types and forms. We also introduce some uncertainty theories, especially DST and its application to confidence propagation. Finally, we review some so-called quantitative approaches which use DST to assess confidence in argument structures.

In chapter 2, we present an approach to confidence propagation in GSN-modeled argument structure based on previous works reviewed in chapter 1. Our approach, uses propositional logic to express the logical relation between a conclusion and the premise(s) supporting it for each argument type. Next, it uses DST to quantify confidence in this support relation along with premises to calculate the overall confidence in the conclusion. Then, we compare our argument types with those reviewed in chapter 1. We also conduct a sensitivity analysis on these argument type to study their behaviours.

In chapter 3, we present an expert opinion elicitation approach, that collects assessments on argument structures and transform them into an appropriate form in the setting of DST (belief degrees). We also justify the use of the "*pignistic*" transform, introduced in chapter 1 and used in our elicitation model, by comparing it with another transform.

In chapter 4, we present the overall confidence/uncertainty assessment and propagation methodology in GSN-based argument structures. It uses the elicitation

model (presented in chapter 3) to provide inputs to the propagation model (presented in chapter 2), and calculate the overall assessment in the top goal of the argument. Then, we present an application to a real system to validate this method. Finally, we discuss the limitations of this approach in view of the case study results and the sensitivity analysis conducted in the previous chapters.

In chapter 5, we present a new, purely qualitative method of assessing confidence in argument structures. It uses a qualitative counterpart of belief functions to quantify and propagate confidence in GSN. Thus, we propose new qualitative confidence propagation formulas as well as a new elicitation model adapted to this method. We also attempt a first comparison between qualitative and quantitative confidence analysis methods.

In appendix A, we present a questionnaire for estimating uncertainties in assurance cases. This questionnaire was used to collect assessment about the GSN used as case study in chapter 4.

Chapter 1

General background and State of the art

Contents

1.1 Introduction 5
1.2 Argument structures 6
1.2.1 A definition of argument structures 6
1.2.2 Forms of argument structures 7
1.3 An introduction to uncertainty theories 10
1.3.1 The concept of uncertainty 10
1.3.2 Uncertainty modeling 11
1.4 DST applied to confidence propagation 17
1.4.1 DS theory and logic 17
1.4.2 Graphical representation of belief propagation 18
1.5 Quantitative confidence propagation approaches in argument structure 19
1.5.1 Uncertainty theories in arguments 19
1.5.2 Argument type and logic 19
1.5.3 Expert elicitation 23
1.6 Conclusion 24

1.1 Introduction

This thesis addresses the issue of confidence (or uncertainty, considered as two sides of the same coin) assessment in an argument structure by its estimation and propagation from the premises to the general conclusion. To simplify the understanding and recovery of the formulas proposed in this manuscript, a definition of all the theoretical concepts and tools used in this work is presented in this chapter. Thus, we start in section (1.2) with a definition of argument structure and a presentation of their different variants and the tools used to represent them. Then, in section (1.3), we introduce theories used to model and propagate uncertainty, followed by a presentation of some DST applications in confidence propagation in section (1.4).

Finally, in section (1.5), we present some previous approaches to confidence modeling and propagation in argument structures, and we position our approach among them.

1.2 Argument structures

Argument structures are commonly used to argue that a high-level property of a system is satisfied. For instance, we talk about safety argument structure (or safety case) when the property being considered is safety [Kelly 2004]. However, it could be extended to more general concepts like dependability [Weinstock 2004, Matsuno 2010] (which includes safety, security, reliability, etc.), trust [Górski 2005b, Górski 2005a] (including both safety and security) or assurance (for any property) cases.

1.2.1 A definition of argument structures

Any argument structure includes three essential elements [Rushby 2015]: A *statement* (or a conclusion) which states the property to be satisfied (e.g, safety, security, etc.), *evidence* about the system architecture carried by premises, and an *argument* (also named warrant or inference, to differentiate it from an argument structure which refers to the whole case) stating that the evidence is sufficient to establish the statement. Several definitions of argument structures exist in the literature. They may vary slightly according to the field of use (e.g, automotive, railway) or the application (e.g, to safety, dependability). A common definition could be : "*A clearly represented collection of rational pieces of evidence (e.g, test or simulation results, expert judgments, analysis reports, etc.) that a system reaches a required property for a given application and environment*" [Bishop 2000].

Safety arguments are not often explicitly mentioned in safety standards. Actually, most of safety standards provide a list of required pieces of evidence to produce, to justify the safety of a system. This list is based on required integrity levels. For instance, ISO 26262 [ISO 26262 2011] defines a risk classification system for automotive vehicles known as *Automotive Safety Integrity Level* (ASIL). It defines four levels of integrity from ASIL-A to ASIL-D (D being the highest class). Each level is determined according to a risk calculation (mainly using severity and likelihood). The greater the risk is, the higher the integrity level and, therefore, the more critical are its requirements. The concept of classification in levels of integrity is used in several industries, such as railway (CENELEC 50126/128/129) or machinery (ISO 13849), and with other names like DAL, Design Assurance Levels, in avionics (ED-12/DO-178/DO-254). However, these standards (unlike safety cases) do not define an explicit argument that links the required evidence (the list), to the safety goal (the rationale is not provided), which makes them difficult to apply to new technologies, or in context of use. It is worth mentioning that in the automotive standard ISO 26262 and in the railway standard CENELEC 50126, there is a reference to the concept of safety case or safety argument, but it is mainly as

> For hazards associated with warnings, the assumptions of [7] Section 3.4 associated with the requirement to present a warning when no equipment failure has occurred are carried forward. In particular, with respect to hazard 17 in section 5.7 [4] that for test operation, operating limits will need to be introduced to protect against the hazard, whilst further data is gathered to determine the extent of the problem.

Figure 1.1: An example of a textual argument [Kelly 2004]

a supporting tool to provide the list of evidence. The UK defence standard 00-56 [MoD 00-56 2007] is one of the non-presciptive standards that explicitly calls for an argument-based approach to justify safety.

Hence, the argument-based approach is not fully integrated in the current standards, but it is becoming more popular, as it allows to deal with new technologies or new context of use.

1.2.2 Forms of argument structures

The structure of the argument can be represented in natural language in the form of well organized texts and tables in order to argue the respect of safety (for a safety case) or other properties (security, dependability, etc.) requirements. However, such a representation presents certain problems highlighted by [Kelly 2004]. Indeed, intended to be produced and used in multicultural (global) environment these arguments are often written in English. He points, [Kelly 2004], that not all engineers responsible for producing safety cases write clear and well-structured English, which may lead to misunderstandings. Thus, there is no guarantee that all stakeholders involved share the same understanding of the argument. Also, the use of cross-references can be confusing. The reader can easily get drowned in the text (often long) and lose the thread of the argument's objective. The example in figure 1.1, displays the issue of cross-references in a textual argument. To address these issues, among others, Kelly proposed a Toulmin [Toulmin 1958] inspired graphical language named "*Goal Structuring Notation*" (GSN).

Goal Structuring Notation

Goal structuring notation (GSN) is a graphical notation/language which represents argument structures (i.e., safety and assurance cases) in the form of directed acyclic graphs (directed trees or arborescences). Figure 1.2 displays its main components.

A GSN breaks down a top claim, called *goal*, into elementary sub-goals following a specific *strategy* and in accordance with a particular *context*. The strategy component, always associated with a goal, justifies its decomposition into other goals,

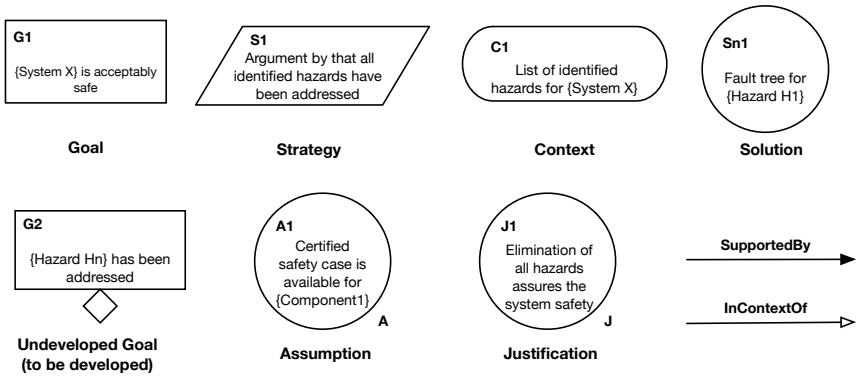


Figure 1.2: GSN main components

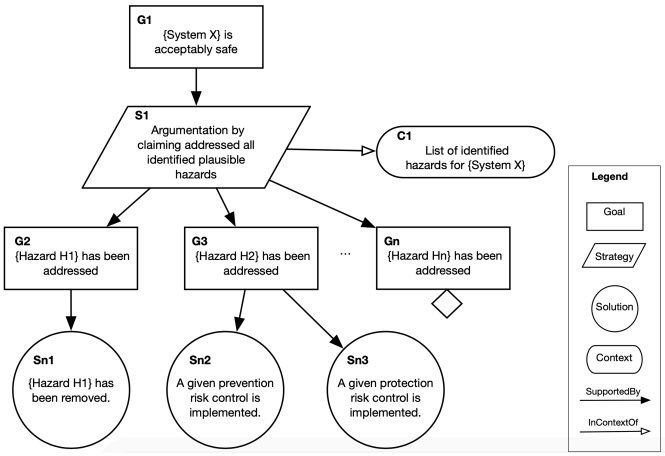


Figure 1.3: GSN example adapted from the Hazard Avoidance Pattern [Kelly 1997]

while the context limits the scope of the argument to the information it provides. This component is mostly related to strategies, but it can be also associated to other components. Each sub-goal is associated with pieces of evidence, called solutions, which support the conclusion. The *assumption* and *justification* components (resp.) bring additional information necessary to (resp.) adopt and justify a strategy. The diamond at the bottom of the goals indicates that they are not yet developed. This representation offers two tools to link components with one another. The *in context of*, represented by a blank tip arrow, links the context component to the others. The *supported by*, represented by black tip arrow, links other components.

Figure 1.3 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 (e.g, Sn_1, Sn_2, ...) following the strategy (S_1).

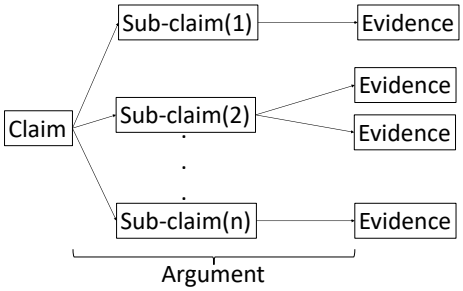


Figure 1.4: An example of a claim-argument-evidence (CAE) notation

Other graphical representations

Beside GSN, other graphical notations exist which indeed have the same purpose, but also use more or less the same concepts/components. For instance CAE (Claim-Argument-evidence) is used to argue the fulfilment of claim (goal in GSN) by proving each of these elements, divided according to an argument (strategy in GSN), by evidence (solution in GSN). Figure 1.4 represents an example of this notation.

Another type of representation known as SACM, standing for Structured Assurance Case Meta-model, was build upon GSN [Kelly 2004] and CAE [Bloomfield 2014] formalism. It is more expressive due to the extensive range of features it offers. For instance, compared to a GSN, it offers more types of link between components (e.g, asserted, need-support, defeated, etc.). Moreover, it offers the possibility of displaying counter-examples which describe (through evidence) the situation where the claim is not fulfilled. [Selviandro 2020] discusses some differences between GSN, CAE and SACM. Different assurance cases modeling tools are available with specific settings for each graphical notation [Wei 2019]. For instance, D-Case Editor [Matsuno 2010] or AdvoCATE [Denney 2018].

In the context of our work, we focus in the following on argument structures modeled with GSN formalism only. Indeed, we do not need all the expressiveness of SACM for instance. Moreover our approach is focusing on the core concepts of arguments and is thus applicable to all representations of arguments.

GSN applications

Categorized as a qualitative approach to logical argumentation [Duan 2014], GSN is used in different industries, especially those which involve critical systems. For instance, [Palin 2010, Habli 2010, Leppäaho 2021] use GSN to argue the safety of automotive systems in accordance with the safety standard ISO26262. On the other hand, [Gallina 2016] uses it to justify safety of rail vehicles.

Machine learning (ML) based systems, among others, also use GSN to justify their safety [Picardi 2020, Gauerhof 2020, Porter 2022]. For example, [Hawkins 2021] proposed a collection of safety cases and processes to be used to establish a justified confidence in autonomous ML-based systems. Authors in [Damour 2021], used a formalism close to GSN to prove the safety of a hybrid

architecture of an anti-collision system based on neural network.

GSN is becoming popular in the safety case field and is even referenced in the automotive safety standard ISO 26262 [ISO 26262 2011].

Discussion

Graphical notations, and GSN in particular, address the concerns raised above by offering a clear, relatively short and well-structured representation, making it more expressive and thus easier to comprehend. However, this symbol-based language does not capture the uncertainty that may exist in the argument structure. This uncertainty may be caused by the lack of specification of the nature of the logical links between the goals in their support to the top one (e.g, link between $G_1, G_2, \ldots G_n$, in figure 1.3). Thus, one cannot properly formalize the uncertainty propagation scheme. The uncertainty related to these links could be interpreted as an indicator of the degree of confidence in the chosen strategy (S_1, justifying the decomposition into sub-goals). In the same way, a GSN does not indicate the confidence that can be associated with the evidence (Sn_1, Sn_2 and Sn_3). This issue was highlighted in several works ([Bloomfield 2010], for example). So-called quantitative approaches have been proposed to overcome these problems. Works of [Duan 2014, Graydon 2017] stated and discussed some propositions that deal with the issue of uncertainty.

1.3 An introduction to uncertainty theories

In order to deal with sources of uncertainty in GSN base assurance cases, this section presents some theoretical background about the concept of uncertainty and presents some tools to model and propagate it.

1.3.1 The concept of uncertainty

In the following, we define the concept of uncertainty and present some classifications of its sources.

Uncertainty definition

The term uncertainty refers to lack of certainty about the truth or falsity in a piece of information (or a proposition). Information is either, *objective* (e.g, measurements produced by a sensor) or *subjective* (e.g, observations or testimonies produced by individuals). The attribute could also take a *quantitative* format (e.g, temperature measurement or an age of a person) or a *qualitative* one (e.g, color or shape of an object) expressed with a natural language. Finally, information could be *singular* produced from particular event in the case of a testimony or an observation; or *generic* resulting from a collection of situations (e.g, scientific laws, commonsense knowledge or statistical models). This variation in types of information leads to

several sources of uncertainty which can be represented either in quantitative or qualitative terms.

Sources of uncertainty

There are as many classifications of uncertainty as there are of information. The most common one distinguishes between: *Aleatory* uncertainty derived from objective or generic information and due to randomness of natural phenomena (e.g., rainfall in the Tropics or snowfall in Himalaya) or in games (e.g., dice rolling) measured by frequency. And *epistemic* uncertainty derived from subjective or singular information and due to the lack of knowledge, e.g., "a crime suspect fled in a grey car". Notice that this information is not that sufficient to track down the suspect. What kind of car was it? In which direction did he/she flee?

A second classification introduces a third category of uncertainty. In addition to the uncertainty due to *randomness* (aleatory) and *incompleteness* (epistemic), another source due to *inconsistency* was added. It represents contradiction (or conflict) between pieces of evidence when too much information is available. Unraveling what is true from what is false seems then difficult and confusing (e.g, pro- and anti-vaccine arguments/researches in a global pandemic situation).

A third classification introduced sources related to *fuzziness* and *vagueness*. These two concepts are often confused. Their meanings are different. Vagueness represents uncertainty of meaning. For example, TALL is a Boolean concept but we do not know the limit between Tall and not Tall. While fuzziness is a concept that applies to a degree. I know the height of John, 1.70m, but he is tall only to a certain extent.

Another classification [Harmanec 1999] considers two sources of uncertainty. Uncertainty due to *non specificity* which describes the inability of inferring the truth from multiple possible scenarios, and another due to *conflict* which describes inconsistency and disagreement in information.

1.3.2 Uncertainty modeling

A general representation of set of states of affaires viewed as a proposition can be obtained using a function $g(A) \in [0, 1]$, that quantifies the confidence in a proposition A (which can be true or false) such that:

$$\begin{cases} g(\emptyset) = 0,\ g(\Omega) = 1 \\ A \subseteq B \Rightarrow g(A) \leq g(B) \text{ (monotonicity with respect to inclusion)} \end{cases} \tag{1.1}$$

Where $\emptyset$ represents the empty set ($g(\emptyset)$ is the confidence in the impossible event) and Ω the universe of all possibilities ($g(\Omega)$ is the confidence in the fact that the sure event happens).

Uncertainty, (i.e., lack of confidence), and confidence are inversely proportional to each other. High confidence in a statement (or in its negation) means low uncertainty about it, and vice versa.

The function g is a probability function P, if it satisfies the additivity property: $P(A \cup B) = P(A) + P(B)$, if $A \cap B = \emptyset$. More functions, can be derived under some assumptions. For instance, belief [Shafer 1976] (necessity, certainty) or plausibility (possibility [Dubois 1988, Zadeh 1999]) functions. Despite the few differences that may exist between these theories, each of them is designed to deal with specific situations. For instance, possibility functions can deal with uncertainty represented with qualitative values. However, they all serve the same purpose, which is to quantify uncertainty and deal with it through the different tools (operations) they offer. They are often used in applications as pairs such as (belief, disbelief), (belief, plausibility) or (necessity, possibility).

Probability theory

It is known that probability properly deals with aleatory uncertainty (frequentist approach), but fails to represent the epistemic one to some extent. To represent total uncertainty, all elements ω_i of the universe of possibility (Ω) get the same probability value $P(\omega_i) = \frac{1}{|\Omega|}, \omega_i \subseteq \Omega$. For instance, if a state of light in a room is unknown, it is represented with $P(light\ ON) = P(light\ OFF) = \frac{1}{2}$. Without any knowledge, there is no reason to assign a different probability to one outcome rather than to another. The same distribution (0.5, 0.5) is actually used when a lot of information is provided equally for ON and OFF. Thus, with probability it is not possible to make the difference between a completely unknown situation, and a well known random situation. This vision is highly debated among uncertainty specialists. Other counter-examples are proposed in the literature. For example, we can classify the world outside the galaxy into two categories: there are living beings (L) or not (N_L). Hence, the probability of getting one or another is $P_1(L) = P_1(N_L) = \frac{1}{2}$. We can also decompose the living into incompatible events, for example, vegetable only (V) and animal (A) entities. So, the probability of getting one of these three entities is $P_2(V) = P_2(A) = P_2(N_L) = \frac{1}{3}$. Since living is a disjunction of vegetable and animal, the distributions P_1 and P_2 are not consistent: $P_2(L) = P_2(V \cup A) = P_2(V) + P_2(A) = \frac{2}{3} \neq P_1(L) = \frac{1}{2}$.

These drawbacks affect all probabilistic representations of subjective belief in case of incomplete information, in particular Bayesian networks, which are nevertheless relevant for frequentist representations based on sufficient amounts of experimental data. This approach is widely used in the context of confidence propagation. But as illustrated in the paragraph before, it has limitations when representing epistemic uncertainties. Hence there is a need to exploit new theories like DST that are more appropriate to model incomplete information.

Dempster-Shafer theory

As a generalization of probability theory, Dempster-Shafer theory [Shafer 1976] (DST, known also as evidence theory) offers tools to model both aleatory and epistemic uncertainty.

A mass function, or basic belief assignment (BBA), is a probability distribution over the power set of a set Ω, known as the *frame of discernment* (universe of possibilities). Formally, a mass function $m : 2^\Omega \to [0, 1]$ is such that:

$$\sum_{E \subseteq \Omega} m(E) = 1,\ m(\emptyset) = 0 \tag{1.2}$$

Any subset E of Ω such as $m(E) > 0$ 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. Hence this approach handles incomplete information in an unbiased way.

A mass assignment induces a so-called belief function $Bel : 2^\Omega \to [0, 1]$, defined by:

$$Bel(A) = \sum_{E \subseteq A,\ E \neq \emptyset} m(E) \tag{1.3}$$

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: $Disb(A) = Bel(\neg A)$ called disbelief; the value $Uncer(A) = 1 - Bel(A) - Disb(A)$ quantifies the lack of information about A.

Mass functions also induce a so-called plausibility function $Pl : 2^\Omega \to [0, 1]$, defined by:

$$Pl(A) = \sum_{E \cap A \neq \emptyset} m(E) \tag{1.4}$$

Belief and plausibility functions are related: $Pl(A) = Bel(A) + Uncer(A) = 1 - Disb(A)$. They respectively represent the lower and upper limits of probability: $Bel(A) \leq P(A) \leq Pl(A)$. To get uncertainty in a proposition A, two pieces of information are needed: ($Bel(A)$, $Pl(A)$) which gives direct information about the uncertainty range or ($Bel(A)$, $Disb(A)$) which gives direct information about the direction in which the truth is most likely to be: A, if $Bel(A) > Disb(A)$ or $\neg A$ if $Disb(A) > Bel(A)$.

Here are some particular cases of mass functions:

- When focal sets are singletons $E_i = \{\omega_i\}$, with $\Omega = \{\omega_1, \ldots \omega_n\}$: both the plausibility and belief of an event A are equal to its probability: $Bel(A) = P(A) = Pl(A)$, because $E_i \subseteq A$ is equivalent to $E_i \cap A \neq \emptyset$ and $\omega_i \in A$.

- When focal sets are nested $E_1 \subset E_2 \subset \ldots E_n$: Firstly, the plausibility of the union of two events A and B is equal to the maximum plausibility of the individual events. Hence, $Pl(A \cup B) = \max[Pl(A), Pl(B)]$. Thus, plausibility can be seen as a measure of possibility $Pl(A) = \Pi(A)$.

  Secondly, the necessity measure of the intersection of two events A and B is equal to the minimum necessity of the individual events: $N(A \cap B) = \min[N(A), N(B)]$. Thus, necessity can be seen as a measure of belief $N(A) = Bel(A)$.

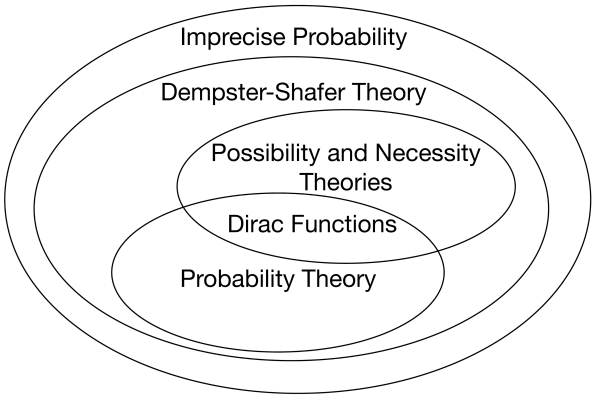


Figure 1.5: Relation between uncertainty theories

- A simple support function [Shafer 1976] shares a mass between a single subset E of Ω (the frame of discernment) and Ω itself. It is relevant to model a testimony from an unreliable source, where we assign a probability s to express that the testimony is irrelevant, so that $m(E) = 1 - s$.

Figure 1.5 represents the relation between different uncertainty theories. We can notice that imprecise probabilities represent a general frame for DST, while both possibility-necessity and probability measures represent special cases. Imprecise probability theory defines a lower $P_*(A)$ and upper $P^*(A)$ probabilities [Walley 1991] that restricts the probability of an event $P(A)$. The width of the interval $[P_*(A), P^*(A)]$ represents the degree of ignorance that an agent has about a proposition A. When the latter is equal to the unit interval $[0, 1]$, we can consider that the agent has no information. In DST these upper and lower probabilities are respectively plausibility and belief degrees.

DST offers tools for dealing with uncertainty, mainly those that allow combination of pieces of information, see [Sentz 2002, Smets 2007] for surveys. Here are some operations that are useful in the context of our work:

- **Discounting:** used to mitigate information (testimonies) coming from unreliable sources. We distrust mass m_Ω with probability v.

$$m^*_\Omega(A) = \begin{cases} v \cdot m_\Omega(A), \text{ if } A \neq \Omega \\ 1 - v \cdot [1 - m_\Omega(A)], \text{ if } A = \Omega \end{cases} \tag{1.5}$$

- **Extension:** used to manipulate two pieces of information that belong to different frames of discernment (Ω_1, Ω_2) by extending it to a common one ($\Omega = \Omega_1 \times \Omega_2$) without further information.

$$m_{\Omega_1 \uparrow \Omega}(A) = \begin{cases} m_\Omega(B), \text{ if } A = B \times \Omega_2, \text{ for some } B \subset \Omega_1 \\ 0, \text{ otherwise} \end{cases} \tag{1.6}$$

Where : $\Omega_1 \uparrow \Omega$ denotes the cylindrical extension of Ω_1 to Ω.

- **Projection:** used to recover a piece of information from a general frame of discernment ($\Omega = \Omega_1 \times \Omega_2$) to its original one ($\Omega_1$).

$$m_{\Omega\downarrow\Omega_1}(A) = \sum_{B\subset\Omega, B\downarrow\Omega_2=A} m_\Omega(B), \forall A \subset \Omega_2 \tag{1.7}$$

Where : $\Omega_1 \downarrow \Omega$ denotes the projection (known also as marginalization) of Ω to Ω_1.

- **Conjunctive rule of combination:** 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. It is $m_\cap = m_1 \otimes m_2$ for $n = 2$, such that:

$$m_\cap(A) = \sum_{E_1\cap E_2=A,\ A\neq\emptyset} m_1(E_1) \cdot m_2(E_2) \tag{1.8}$$

In DST, an additional step eliminates conflict that may exist by means of a normalization factor (dividing $m_\cap$ by $1 - m_\cap(\emptyset)$). This is Dempster rule of combination [Shafer 1976]. The mass of the void $m_\cap(\emptyset)$ represents the conflict degree. When it is small, this normalisation is meaningful: it eliminates the conflict and proportionally increases the contradiction-free degrees $m_\cap(A)$. On the contrary if conflict is important (near 1), normalisation loses sense (division by 0). A possible alternative is to consider the conflict as ignorance, hence we can add $m_\cap(\emptyset)$ to the mass of the tautology $m(\Omega)$ (Yager's rule [Yager 1987]).

Example 1. Let's say that a group of ornithologists (bird experts) want to conduct a study on falcons living in the "Occitanie, France" region. This region is characterized mainly by the presence of two species *Peregrine (P)* and *Kestrel (K)* falcons respectively recognisable by their gray and brown color. $\Omega_1 = \{F, \neg F\}$ refers to the bird species $m_{\Omega_1}(F)$ and $m_{\Omega_1}(\neg F)$ represent evidence respectively supporting that the seen bird is a falcon or not. $\Omega_2 = \{Gr, Br, Bl\}$ refers to the bird color respectively grey (Gr), brown (Br) and black (Bl) seen during this study. We assume that no black falcon specie lives in the studied region. So we have that $P = F \wedge Br$, $K = F \wedge Gr$, where: $P, K \in \Omega$ and $F \wedge Bl = \emptyset$. Each expert testimony is attached a mass function on $\Omega = \Omega_1 \times \Omega_2$.

Lets suppose that an expert thinks that a group of birds are Peregrines falcons with a belief of $m_\Omega(P) = 0.5$ or Kestrel with a belief of $m_\Omega(K) = 0.3$. Hence, we can deduce that his/her belief that the seen birds were falcons is (using the projection operation):

$$\begin{aligned} m_{\Omega_1}(F) &= m_{\Omega\downarrow\Omega_1}(F) \\ &= m_\Omega(P) + m_\Omega(K) \\ &= 0.5 + 0.3 = 0.8 \end{aligned}$$

Now, we suppose that another expert gives the following testimonies about his/her observation.

- $m_{\Omega_1}(F) = 0.4$, $m_{\Omega_1}(\neg F) = 0.2$ and $m_{\Omega_1}(\Omega_1) = 0.4$
- $m_{\Omega_2}(Br) = 0.3$, $m_{\Omega_2}(Gr) = 0.1$, $m_{\Omega_2}(Bl) = 0.4$ and $m_{\Omega_2}(\Omega_2) = 0.2$

Let us also consider that because of bad weather the experts observations may be subject to reasonable doubts. Assuming that testimonies about the species could only be trusted at 0.75%. Thus, the updated masses are (using the discounting operation): $m_{\Omega_1}(F) = 0.75 \times 0.4 = 0.3$, $m_{\Omega_1}(\neg F) = 0.75 \times 0.2 = 0.15$ and $m_{\Omega_1}(\top) = 1 - 0.75 \times (1 - 0.4) = 0.55$. We can notice that due to this additional source of uncertainty the belief in these testimonies decreases, while the uncertainty on them increases. The same operation is applied to m_{Ω_2}.

In order to illustrate how a DS combination rule works, we combine two mass functions defined on Ω_1 and Ω_2, respectively. In fact we combine their extensions on Ω. Table 1.1 presents all possible intersections of focal sets in the two frames of discernment (Ω_1 and Ω_2). For instance $F \wedge Br$ stands for $(\{F\} \times \Omega_2) \cap (\Omega_1 \times \{Br\})$.

Table 1.1: DS combination rule example

$m_\Omega = m_{\Omega_2} \otimes m_{\Omega_1}$	$m_{\Omega_1}(F) = 0.3$	$m_{\Omega_1}(\neg F) = 0.15$	$m_{\Omega_1}(\Omega_1) = 0.55$
$m_{\Omega_2}(Br) = 0.225$	$P = F \wedge Br$	$\neg F \wedge Br$	Br
$m_{\Omega_2}(Gr) = 0.075$	$K = F \wedge Gr$	$\neg F \wedge Gr$	Gr
$m_{\Omega_2}(Bl) = 0.3$	$\emptyset$	$\neg F \wedge Bl$	Bl
$m_{\Omega_2}(\Omega_2) = 0.4$	F	$\neg F$	Ω

We can notice that we have an empty intersection which represents contradiction (since there is no black falcon specie that lives in the studied region). Hence, $m_\Omega(\emptyset) = m_{\Omega_1}(F) \times m_{\Omega_2}(Bl) = 0.3 \times 0.3 = 0.09$.

Let us calculate the belief that the birds observed are not falcons and that they have a brown color (the focal $\neg F \wedge Br$ in grey in the table 1.1). If we do not ignore the effect of the conflict on the value of the obtained belief, we obtain: $m_\Omega(\neg F \wedge Br) = m_{\Omega_1}(\neg F) \times m_{\Omega_2}(P) = 0.15 \times 0.225 = 0.03375$. As seen, this combination requires normalization dividing by $1 - m_\Omega(\emptyset)$ to eliminate conflict. Hence, $m_\Omega(\neg F \wedge Br) = \frac{m_\Omega(\neg F \wedge Br)}{1 - m_\Omega(\emptyset)} \approx 0.0371$. We can notice that this normalization increases the mass of $m_\Omega(\neg F \wedge Br)$.

Examples of other combination rules

There are also several variants of Dempster rule of combination that could be used in evidence fusion (e.g, Smets rule [Smets 2005] similar to the DS combination rule but without normalisation, Inagaki's [Inagaki 1991], Yager's rule [Yager 1987] or the weight average rule [Sentz 2002]). However, each method obeys certain assumptions and describes some kind of situation. This is why it is necessary to choose the most appropriate combination rule for our 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 [Denoeux 2008, Destercke 2011].

Transformation from belief to probability

As pointed out above, many theories that deal with uncertainty exist. They either are a particular case of a general one (e.g, probability and DS Theory) or have common bases (e.g, Fuzzy sets and DST). Figure 1.5 illustrates this intersection. It is therefore interesting to see how one can switch, if possible, from one uncertainty measure to another.

In this paragraph we present two common transformation of mass (BBA) to probability that will be used in the next chapters :

- **Shenoy transform [Cobb 2006]:** This transformation divides the plausibility of a singleton $Pl(\omega)$ (with $\omega \in \Omega$), by the sum of plausibility of all elements $Pl(\omega')$ (with $\omega' \in \Omega$). But, it is inconsistent with uncertainty interval, namely we may have $P_{Sh}(A) \notin [Bel(A), Pl(A)]$, where:

$$P_{Sh}(\omega) = \frac{Pl(\omega)}{\sum_{\omega' \in \Omega} Pl(\omega')} \tag{1.9}$$

- **Pignistic probability transform [Smets 2005]:** This transformation turns a mass m on a set Ω into a probability, changing the focal sets into uniform distributions. It represents a generalization of Laplace insufficient reason principle and coincides with *Shapley value* in game theory.

$$BetP(\omega) = \sum_{E:\omega \in E} \frac{m(E)}{|E|} \tag{1.10}$$

1.4 DST applied to confidence propagation

This paragraph presents and discusses some applications of uncertainty theories, specially DST, to confidence propagation.

1.4.1 DS theory and logic

Logical reasoning and belief functions are not often put together. An approach to reasoning with Dempster rule of combination was proposed by [Chatalic 1987a]. In this approach each formula in a knowledge base is viewed as a simple support function and combined with other formulas in the knowledge base. In other works, [Banerjee 2014, Cholvy 2015, de Saint-Cyr 1994] explored the semantic/interpretation side of belief function combination and its relation to modal logic. DST can be also used to combine logic and probability as proposed by [Cozman 2019].

In relation with our work, Wang [Wang 2016, Wang 2019] uses logic to describe the support of premises to a conclusion and DST to quantify confidence in that support. This approach consider two sets (resp.) of n Boolean variables $X_i \in \{x_i, \neg x_i\}$ and well-formed formulas $KB = \{\phi_1, \ldots \phi_n\}$ in propositional logic (knowledge base), and a formula C representing a conclusion that can be inferred

from the knowledge base, such that $KB \vdash C$. Assume each formula ϕ_i (a combination of X_i and $\{\neg, \wedge, \vee\}$ connectors) 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 $\top$ summing to 1. Take for example the case of a simple premise P and a rule in the form of an equivalence $P \equiv C$ as in Wang. 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 $(m_2(P \equiv C) + m_2(\top) = 1)$. Using an appropriate combination formula in the setting of DST the belief in the conclusion C based on beliefs about premises and rules ϕ_i is computed. Extending classical logic inference to this uncertain environment can be done by means of Dempster rule of combination [Chatalic 1987b], 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 $Bel(C) = \sum_{\phi_i \vdash C} m(\phi_i)$.

1.4.2 Graphical representation of belief propagation

Like Bayesian networks, there are also graphical representations of beliefs. For instance, *valuation networks* are graphical representation of valuation-based system (VBS) introduced by [Shenoy 1989] used to propagate uncertainty using belief functions. As defined by [Jiroušek 2014], a VBS is a graphical tool that allows, in its static part, to represent knowledge through a finite set Φ of variables $\Phi = \{D, G, B, \dots\}$. The knowledge encoded on a subset of variables, modeled by BPA (or masses) when DST belief functions are used, is called valuation $\Psi = \{\delta, \gamma, \beta, \dots\}$. Any subset of Φ is represented by a lower-case Roman alphabets (e.g, $r, s, t, \dots$). For instance, δ is called a valuation for r, where $r \subseteq \Phi$. It has also a dynamic part that concerns the reasoning with knowledge which includes three main operations (i.e, marginalization, combination and removal). These operations are used to compute uncertainty.

Figure 1.6 represents an example of a valuation network, where circles represents variables and the diamonds represent the valuations related to each of these variables. We notice that the semantics of an argument does not exist in this representation and that the link between the variables is not explicit which discourages its use for confidence assessment in an argument structure.

Uncertainty propagation can be addressed by standard existing belief function software based on results in [Shenoy 1990] (e.g., the belief function machine implemented in MatLab), but the GSNs we study have a particular tree-like structure that enables an explicit symbolic calculation of the belief function on the conclusion space. Explicit formulas, that can be obtained from approaches such as the one we propose, make the calculation more efficient and we can predict the effect of changing selected inputs, thus better explaining the obtained results, and validating the approach.

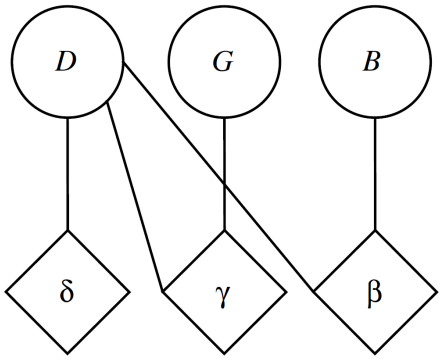


Figure 1.6: An example of a valuation network

1.5 Quantitative confidence propagation approaches in argument structure

In this section, we present main research works which focus on expressing confidence in the safety case, using uncertainty theories.

1.5.1 Uncertainty theories in arguments

Several research works use probability theory to model uncertainty and propagate it with Bayesian networks [Denney 2011, Hobbs 2012, Guiochet 2014]. For instance, some authors [Nešić 2021] transform GSN into a Bayesian network (BBN) and propagate probabilities accordingly. Due to the limited expressiveness of the probabilistic framework, 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 a lot of time in order to collect and process it.

Other less common approaches are used for confidence propagation. For instance, an approach based on Baconian probabilities (a variant of necessity measure) was proposed to identify and eliminate sources of uncertainty (defeaters) in argument structures [Weinstock 2013]. Another one used subjective probability to assess confidence in argument structures [Yuan 2017].

1.5.2 Argument type and logic

Managing uncertainty more effectively, several works used DST to model and propagate confidence in GSN. They usually proposed different propagation formulas for different identified types of arguments [Ayoub 2013, Cyra 2011, Wang 2019]. However, they poorly describe how premises interact to support the conclusion (argument types). For instance, Anaheed et al. [Ayoub 2013] explain these interaction using Venn diagrams (see figure 1.7) which actually mixes the logical (conjunction, disjunction, etc) and confidence (mass assignment) aspects of these types. This approach is misleading. They use weighted average to describe complementarity

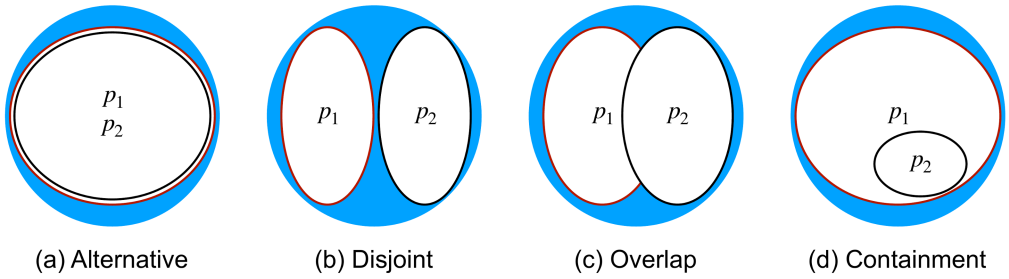


Figure 1.7: Argument type proposed by Annaheed et al. [Ayoub 2013]

(*disjoint argument*, all premises are needed to support the conclusion), while they use Dempster combination rule to represents redundancy (*alternative argument*, each premise can support the conclusion alone). In situations, where there is a combination between these two (alternative and disjoint) they also use a combination between Dempster rule and weighted average. For example, the formula below is used to calculate the confidence in the conclusion (C) supported by two premises (p_1 and p_2) for the *overlap argument*.

$$Bel(C) = \frac{\omega_1 \times Bel_1(p_1) + \omega_2 \times Bel_2(p_2) + \omega_{Overlap} \times [Bel_1(p_1) \otimes Bel_2(p_2)]}{\omega_1 + \omega_2 + \omega_{Overlap}}$$

Where: ω_1, ω_2 and $\omega_{Overlap}$ quantify respectively the degree of contribution of premises p_1, p_2 and their overlap. These values are not confidence degree ($\omega_i \in \mathbb{N}^*$). This calculation supposes that the argument can be divided into disjoint and alternative sub-arguments. For example, p_1 justifies functions from 1 to 4, and p_2 functions 2 to 7. Thus, we can assign the following values to ω_i: $\omega_1 = 4$, $\omega_2 = 6$ and $\omega_{Overlap} = 3$. The overlapping parts are considered as an alternative argument type. After combination (using Dempster combination rule), the argument is considered then as an disjoint argument with the remaining parts (to combine using weighted average). We agree that their proposal is an argument type that expresses a combination between complementarity (conjunction) and redundancy (disjunction). However, we do not believe that this argument can always be broken down into a complementary and a redundant part. Let us consider that a system is acceptably safe (C) because the test results (p_1) and formal verification (p_2) are satisfactory. This example represents a typical case of an overlapping argument, which we cannot divide it into two sub-arguments (alternative and disjoint), and therefore, the corresponding formula cannot be applied. Moreover, this work does not propose formulas for the propagation of disbelief in the conclusion ($Disb(C)$) coming from the premises as defined in the DST setting ($Bel, Disb, Uncer$). This information can be useful to differentiate the case where the system is not safe (low value of $Bel(C)$) because of lack of information (high value of $Uncer(C)$) or because the premises are not trustworthy (high value of $Disb(C)$).

In Cyra and Gorski's works [Cyra 2011] the pair (belief,plausibility) is used to

Table 1.2: Propagation formulas from [Cyra 2011]

Argument type	formulas
C-Argument	$\begin{cases} Bel(C) = Bel(W) \times \frac{k_1 \times Bel_1(p_1) + k_2 \times Bel_2(p_2)}{k_1 + k_2} \\ Pl(C) = 1 - Bel(W) \times (1 - \frac{k_1 \times Pl_1(p_1) + k_2 \times Pl_2(p_2)}{k_1 + k_2}) \end{cases}$
NSC-argument	$\begin{cases} Bel(C) = Bel(W) \times Bel_1(p_1) \times Bel_2(p_2) \\ Pl(C) = 1 - Bel(W) \times [1 - Pl_1(p_1) \times Pl_2(p_2)] \end{cases}$
SC-argument	$\begin{cases} Bel(C) = Bel(W) \times Bel_1(p_1) \times Bel_2(p_2) \\ Pl(C) = 1 \end{cases}$

quantify uncertainty in argument structures. They also use natural language to describe their argument types, which make it difficult to identify (from the formulas) the rationale that leads to them. Thus, we cannot reproduce all of them. To illustrate this issue, we present some propagation formulas (see table 1.2) to discuss. These formulas calculate the confidence in a conclusion (C) supported by two premises (p_1 and p_2). In this table, C-Argument stands for Complementary argument, NSC for Necessary and Sufficient Condition list argument and SC for Sufficient Condition list argument. $Bel(W)$ quantifies the belief in the strategy, and k_i represents the weight of each premise. We cannot determine whether k_i is a coefficient (as in the case of Anaheed et al.) or a degree of confidence (as in Wang et al.). Note that all argument types defined in Cyra and Gorski's work (NSC, SC, C arguments, etc.) will be discussed in detail later in chapter 2. Saying that a conclusion (C) is supported by complementary (C-Argument) premises means that they are all needed to infer the conclusion. However, Cyra and Gorski present three different formulas the describe the same notion. C-Argument uses a weighted average (same as the disjoint argument with Anaheed et al.) attenuated by the belief on the strategy $Bel(W)$. While NSC argument uses the product of premises belief and a sort of probabilistic sum of disbelief ($Disb(C) = 1 - Bel(W) \times [1 - (1 - Disb(p_1)) \times (1 - Disb(p_2)]$) both attenuated by the belief on the strategy. The SC-argument represent a particular case of the NSC-argument type where $Disb(C) = 0$ ($Pl(C) = 1$). Thus, one can question the interest of defining such an argument.

In Wang et al. works [Wang 2018a, Wang 2018b] partially based on Cyra and Gorski, which are close to ours, a separation between the logic (argument type definition) and confidence (mass assignment) was made, which facilitates the reproduction of their formulas. They use a mix between conjunction and disjunction to model their types and logical equivalence to link premises to the conclusion. In this work, they actually use logical equivalence to express inference (logical inferences will be defined and named "rules" later in chapter 2) between premises and the conclusion. Using equivalence for inference, means that we assume that true premises lead to a true conclusion. Conversely, if the premises are false, the conclusion is also false. However, in an argument, false premises does not necessarily means that the conclusion is false. Moreover, such confidence propagation models usually rely

Table 1.3: Propagation formulas from [Wang 2019]

Argument	formulas
C-Arg	$\begin{cases} Bel(C) & = \omega_1 \cdot Bel_1(p_1) + \omega_2 \cdot Bel_2(p_2) \\ & +(1-\omega_1-\omega_2) \cdot Bel_1(p_1) \cdot Bel_2(p_2) \\ Disb(C) & = \omega_1 \cdot Disb_1(p_1) + \omega_2 \cdot Disb_2(p_2) \\ & +[1-(1-\omega_1-\omega_2) \cdot (1-Disb_1(p_1)) \cdot (1-Disb_2(p_2))] \end{cases}$
R-Arg	$\begin{cases} Bel(C) & = \omega_1 \cdot Bel_1(p_1) + \omega_2 \cdot Bel_2(p_2) \\ & +[1-(1-\omega_1-\omega_2) \cdot (1-Bel_1(p_1)) \cdot (1-Bel_2(p_2))] \\ Disb(C) & = \omega_1 \cdot Disb_1(p_1) + \omega_2 \cdot Disb_2(p_2) \\ & +(1-\omega_1-\omega_2) \cdot Disb_1(p_1) \cdot Disb_2(p_2) \end{cases}$

on experts to provide the belief degrees of the premises necessary to run them (this issue will be detailed in the next subsection). In practice, one may find a situation where an expert can only judge whether the conclusion is true or not and nothing else. The models presented in this work do not consider this case where a single type of information on the conclusion (for or against) is provided.

On the other hand, in Wang et al. they assign all inference expressions (ω_i), defining the argument types, to a single mass function ($\sum \omega_i = 1$). For instance, $m(\omega_i)$ in the case of the complementary argument type, we get the following inference expressions (for an argument with two premises): $m(p_1 \equiv C) + m(p_2 \equiv C) + m([p_1 \wedge p_2] \equiv C) + m(\top) = 1$. The first and second terms respectively quantify the support coming from the premises p_1 and p_2, while the third term ($m([p_1 \vee p_2] \equiv C)$ for the redundant argument type) quantifies the support of the conjunction (resp, disjunction) of the two premises. The last term (mass on the tautology) quantifies the amount of ignorance. This choice makes the confidence in one premise dependent on confidence in the others. The more confidence in one premise is high, the less confidence in the others should be. Having one mass function for each logical formula allows the assessment of each premise independently from the others and reveal the presence of conflicts where we have opposite assessment about different premises supporting the same conclusion. Table 1.3 presents propagation formulas of the complementary (C-Arg) and redundant (R-Arg) argument types. Where: $\omega_i = m(p_i \equiv C)$ and, $1 - \sum_{i=1}^{n} \omega_i = m([p_1 \wedge p_2] \equiv C)$ for the C-Arg or $m([p_1 \vee p_2] \equiv C)$ for the R-Arg ($n = 2$, the number of premises). They also define three other argument types which can be considered as particular case of these two: Full complementary (FC-Arg), full redundant (FR-Arg) and disparate argument type (D-Arg). Notice that that all formulas take the form of a sum of products. This is due to the use of a single mass function for all rules. Furthermore, we notice that the belief formulas of C-Arg and the disbelief one of R-Arg take the same pattern. The same is true for the two remaining formulas. For example, when $\omega_1 = \omega_2 = 0$, the belief formula for C-Arg takes a conjunctive form ($Bel(C) = Bel_1(p_1) \cdot Bel_2(p_2)$) and disbelief take a disjunctive one (probabilistic sum: $Disb(C) = 1 - [1 - Disb_1(p_1)] \cdot [1 - Disb_2(p_2)]$). Similarly for the R-Arg. This results is consistent with the definition of the argument types. Thus, unlike

the formulas proposed in Cyra and Gorski, and Annaheed et al. works (mainly those that use the weighted average) one can always figure out what they express by looking at the forms they take (product for conjunction, and the probabilistic sum for disjunction).

We also notice that none of approaches that we have seen considers the case of conflict (opposite assessments for the same conclusion). In this case the mass of the conflict $m(\emptyset)$ takes a positive value. This mass is calculated from the empty intersection between logical expressions. Yet, in practice, this situation is common.

1.5.3 Expert elicitation

Building such a confidence model relies on input values, usually provided by experts in qualitative form, and transformed into quantitative values. Hence, there is a need of developing an "Expert opinion elicitation" approach. This method is more often used with probabilistic models. For instance, in [De Persis 2019], authors used an expert elicitation procedure for a risk assessment approach in fault trees. However, it can also be used in evidence theory. Ben Yaghlane et al. [Yaghlane 2008], generate belief functions from a preference relation between events provided by experts. In relation to our framework, few authors augmented their confidence assessment method by such a data elicitation procedure in order to provide quantitative values for their models. Only some authors such as [Cyra 2011, Nair 2015, Wang 2018a] used an elicitation method that transforms expert opinions, about premises, given in the form of qualitative values, into quantitative ones.

Cyra and Gorski proposed a formulas that turn numerical values associated to the decision (in qualitative format) about a premise (between "acceptance" and "rejection") and confidence degree associated to it (from "lack of confidence" to "for sure") into belief, disbelief and uncertainty values, which were reused in Wang et al. works. See the formulas below:

$$Conf = Bel + Disb, \quad Conf \in [0, 1]$$

$$\begin{cases} Dec = \frac{Bel}{Bel+Disb}, & \text{if } Conf \neq 0 \\ Dec = 1 \ \ (0 : \text{in Wang et al. approach}), & \text{if } Conf = 0 \end{cases}$$

This elicitation approach uses a linear scale for both decision and confidence with respectively four and six equidistant items. However, in case of total ignorance ($Conf = 0$) [Cyra 2011] adopted an optimistic point of view by accepting the conclusion ($Dec = 1$), while [Wang 2018b] choose a more conservative approach by rejecting the conclusion ($Dec = 0$). These choices seems arbitrary since based on no evidence. It would be better to take a neutral position because no evidence is provided to justify taking a side. Moreover, to switch back from quantitative values of decision and confidence when $Bel = Disb \neq 0$ (conflict situation), the decision

numerical value takes the midpoint of the interval ($Dec = 1/2$) which makes the choice of the appropriate qualitative decision (whether to accept or reject) very difficult.

Concerning the collect of mass on rules (ω_i), Wang et al. choose to use identification techniques. For that, they ask experts to give their assessments about the conclusion (outputs) according to predefined assessment of premises (inputs). Then, using *Non linear least square method* they identify the values of the rule weight (ω_i). However, we notice that this method could lead to values outside the interval [0, 1], which makes no sense. Moreover, asking the expert to give his assessment of the conclusion according to predefined inputs (i.e, Supposing that we have "*tolerable, with high confidence*" and "*opposable, with low confidence*" assessment on premises, what this your assessment on the conclusion ?) can be disturbing and difficult, specially if you have several premises.

Transforming expert opinions, in the form of decisions and their associated degree of confidence, about the premises into qualitative values of belief and disbelief, and then qualitative values about the conclusion into decision and confidence can also be a source of uncertainty. Indeed, these elicitation approaches numerically encode qualitative values using an arbitrarily chosen equidistant scale.

1.6 Conclusion

In this chapter, we have defined argument structures and pointed out the interest of using graphical formats such as GSN. We have also highlighted the main issue with these representations, namely the inability to capture the uncertainty that may exist in such structures. In order to address this issue, we presented some so-called quantitative approaches compared to the GSN method which is categorized as a qualitative confidence assessment method. These quantitative approaches are built on argument structures to provide models of confidence propagation from premises to an overall conclusion. DST seems today the best option in order to take into account aleatory and epistemic uncertainties, but recent works using DST for argument evaluation have strong limitations regarding their application or justification. Moreover, we have seen that in order to calculate belief degrees, qualitative data is required from experts. This approach is known as Expert Opinion Elicitation. Few works address this issue which is mandatory to have an applicable approach. Hence, we first present in the next chapter a confidence propagation model that addresses the issue raised in works mentioned in this chapter. We will then present in the following chapters :

1. A confidence propagation models that address the issues of argument definition, mass assignment and combination rules mentioned above.

2. An elicitation model which collects and transforms expert opinions about the argument structure and used them to feed our propagation models.

3. A complete confidence assessment procedure that integrates both propagation and elicitation models.

4. A purely qualitative model of confidence propagation which propagates qualitative belief and disbelief values to address the issue of uncertainty caused by qualitative-quantitative transformation.

Chapter 2

Uncertainty propagation model in argument structure

Contents

2.1 Introduction 27
2.2 Confidence measure in an argument structure 28
2.2.1 Formal definition of uncertainty in premises 29
2.2.2 Formal definition of uncertainty in the support relation . . . 29
2.3 Confidence propagation models 30
2.3.1 Simple argument type (S-Arg) 30
2.3.2 Conjunctive argument type (C-Arg) 32
2.3.3 Disjunctive argument type (D-Arg) 35
2.3.4 Hybrid argument type (H-Arg) 37
2.3.5 Conflict mass for (n) premises 43
2.4 Argument types in literature - A comparison 46
2.5 Models sensitivity analysis 48
2.6 Conclusion 52

2.1 Introduction

In arguments like safety cases, the main objective is to take a decision regarding the top goal. For that, we argue that a confidence estimation should be specified. However, this confidence should be propagated from all branches of the argument to the top goal.

Confidence propagation models are built upon an argument structure in order to assess the overall confidence in the top goal. Many works [Ayoub 2013, Cyra 2011, Wang 2019], presented in chapter 1, define several types of argument structure, which describe the relation between premises in support of a top-goal, and propose confidence (belief and disbelief) propagation formulas accordingly. As discussed before, these types are mostly informally defined, and fail to describe this relation (which is essential to define uncertainty propagation formulas). Therefore some authors (e.g, [Graydon 2017]) state that these approaches are not yet sufficiently developed to be applied to real cases.

In the previous chapter, we highlighted the importance of formally defining the argument types, often expressed in natural language, which can be open to various interpretations, in order to avoid any confusion. We also pointed out the benefits of using Dempster-Shafer theory to quantify and propagate confidence/uncertainty in these structures. In the continuity of Wang et al. works [Wang 2018a], we improve this framework by defining more accurate argument types and confidence propagation formulas using resp. logic and DST (since we mostly deal with epistemic uncertainty due to the lack of information). Most of the work has been published in [Idmessaoud 2020, Idmessaoud 2021a].

This chapter is structured as follows. First of all we define the argument types and present the confidence propagation formulas for each one. Then, we compare our argument types with those already proposed in the literature. Finally, we conduct a sensitivity analysis of these types to identify the characteristics and behaviours of each one.

In all that follows, we refer to the top-goal in a GSN as conclusion (C) and the goal(s) supporting it as premise(s) (p_i). The propagation formulas shall be designed on the basis of reduced examples, one (p) or two premises (p_1, p_2) supporting the same conclusion (C) and then generalized for multiple ones (n). These examples will also represent safety cases where the property to verify is safety. However, this approach can be extended to all dependability properties (i.e., security, reliability, etc.) or others.

2.2 Confidence measure in an argument structure

Confidence is opposite to the degree of uncertainty. The higher the measure of confidence in a statement, the lower the value of the uncertainty and vice versa. To properly deal with uncertainty, one should first locate its sources in an argument. As presented in figure 2.1, [Wang 2019] focuses on two sources of uncertainty in GSN based argument structure. The first, related to goals, quantifies the uncertainty in premise(s) "p" and the conclusion "C", commonly known as *trustworthiness*. While the second quantifies the uncertainty in the support between premise(s) and the conclusion, commonly known as *appropriateness*. The later can be related to the *Strategy* component of GSN (see chapter 1), which gives justification of goal decomposition is sub-goals.

The other GSN components can also be a source of uncertainty, like *Assumption* or *Context* components. These components can also be viewed as sources of information that condition the confidence in goals (premises or conclusion). Hence, we can mitigate the confidence related to them with a DST discounting factor (equation 1.5). On the other hand, we can assume that the evaluation given to the premises takes into account all assumptions, justifications and contexts.

In Wang et al. [Wang 2018a] they apply a discounting operation to the conclusion. In DST, this operation is often used when the testimony of an expert is not so reliable. However, since all DST based propagation formulas we previously reviewed

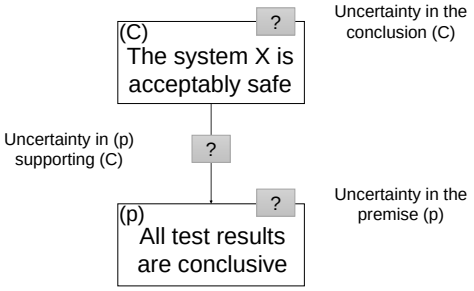


Figure 2.1: Sources of uncertainty in GSN - A simple argument type (S-Arg) example (Note that the arrow direction is not intuitive regarding the implication $p \Rightarrow C$)

in this manuscript rely on the testimonies of experts ([Cyra 2011, Wang 2019]), we can assume that the information comes from highly reliable sources

In follows, we focus only on uncertainties related to premises and conclusions, displayed in figure 2.1.

2.2.1 Formal definition of uncertainty in premises

This parameter is designed to quantify confidence in premises. For instance, in the example of figure 2.1, it answers the question: "*How much these test results are reliable ?*"

It is modeled as a mass function on each premise of the argument, i.e.: $m_p^1, ..., m_p^n$. m_p^i, the i^{th} mass function on the premises, assigns a mass to the premise p_i, one on its negation ($\neg p_i$) and one on the tautology[1] ($\top$, representing ignorance) summing to 1, i.e.: $m_p^i(p_i) + m_p^i(\neg p_i) + m_p^i(\top) = 1$.

2.2.2 Formal definition of uncertainty in the support relation

The support relation between premise(s) and conclusion is translated into a logical expression, that we name a "rule". It designates any statement that can be expressed by a proposition: If "p", then "c" or formally by a material implication: $(p \Rightarrow c) \equiv (\neg p \vee c)$. This rule is designed to quantify the confidence in the claim that the evidence supports the conclusion by evaluating the impact of premises on a conclusion. In the example of figure 2.1, it answer the question: "*Are test results sufficient to guarantee the safety of the system X ?* "

To model it, we associate a simple support function [Shafer 1976] to each rule r^* (defined in the following) of the argument type. Each simple support function consists in assigning a mass $m_r(r^*) = s$ to the rule and another one $m_r(\top) = 1 - s$ to the tautology, these weights summing to 1, i.e.: $m_r(r^*) + m_r(\top) = 1$.

We do not consider here $m(\neg r^*)$, because it supposes that we consider exceptions to the rules where a conclusion is false even if the premises leading to it are true

[1] Also noted $m(\Omega)$ commonly used by safety community. According to the set theory syntax, $[\top] = \Omega$.

or vise-versa. A more detailed explanation will be presented in the next section ("Remarks" paragraph).

2.3 Confidence propagation models

In order to propagate confidence in a GSN pattern, the relation between the premises and the conclusion (which is not formally specified in GSN representations) needs to be modeled first. Logical expressions (i,e. implication, conjunction, disjunction) are chosen for this. We call these logical expression "*rules*". Each one of these rules will be assigned a mass function. These mass functions along with others on premises, will be used to calculate a confidence propagation model using combination tools proposed in DST. Looking through the literature, four common argument types were identified. In the following, we present and explain how confidence is propagated in each case.

2.3.1 Simple argument type (S-Arg)

As indicated in its name, this argument describes the simplest pattern that one can encounter, a conclusion (C) supported by a single premise (p). If the premise is true, then so is the conclusion. Material implication is used to express this rule, called *direct rule* : $p \Rightarrow C$ (which can be also represented with a negation and logical disjunction: $[p \Rightarrow C] \equiv [\neg p \vee C]$). However, implication can only infer the acceptance of the conclusion (when p holds, *Modus ponens* pattern: $\{p,\ p \Rightarrow C\} \vdash C$). If p is false, $\neg C$ cannot be concluded. And since we work in a three-state paradigm : belief, disbelief and uncertainty, an additional rule, called *reverse rule* in opposition to the first one, is used to infer the rejection of the conclusion: $\neg p \Rightarrow \neg C$. Direct and reverse rules are designed (resp.) to propagate belief and disbelief of the premises to the conclusion. This proposal is a first major difference with previous works which only consider $p \Rightarrow C$ or $p \equiv C$. Figure 2.1 presents an example of a simple argument. In this case, the conclusion "The system is acceptably safe" is achieved, only if the premise "The test results are conclusive" is true (direct rule). Otherwise, the conclusion is false (reverse rule). For each rule (direct and reverse), we assign resp. a simple support mass function ($m_{\Rightarrow}$ and $m_{\Leftarrow}$) which puts on each a mass on the rule and another on the tautology ($\top$) summing to 1, i.e.: $m_{\Rightarrow}(p \Rightarrow C) + m_{\Rightarrow}(\top) = 1$ and $m_{\Leftarrow}(\neg p \Rightarrow \neg C) + m_{\Leftarrow}(\top) = 1$. We also assign an other function to the premise (m_p), which puts a mass on the premise (p), its negation ($\neg p$) and the tautology ($\top$) summing to 1, i.e.: $m_p(p)+m_p(\neg p)+m_p(\top) = 1$. Using the conjunctive rule of combination (equation 1.8) to merge the mass on rules (in table 2.1), then its result ($m_r = m_{\Rightarrow} \otimes m_{\Leftarrow}$) with the mass on the premise (m_p) in table 2.2, we obtain the propagation formulas for an S-Arg, in equation 2.1 ($m_C = m_r \otimes m_p$). Some calculation explanations are given after here in the "Remarks" paragraph.

Table 2.1: Combination of direct ($m_{\Rightarrow}$) and reverse ($m_{\Leftarrow}$) rules for S-Arg

$m_r = m_{\Rightarrow} \otimes m_{\Leftarrow}$	$m_{\Rightarrow}(p \Rightarrow C)$	$m_{\Rightarrow}(\top)$
$m_{\Leftarrow}(\neg p \Rightarrow \neg C)$	$p \equiv C$	$\neg p \Rightarrow \neg C$
$m_{\Leftarrow}(\top)$	$p \Rightarrow C$	$\top$

Table 2.2: Combination of the mass on premise (m_p) with its rules (m_r) for S-Arg

$m = m_p \otimes m_r$	$m_r(p \equiv C)$	$m_r(p \Rightarrow C)$	$m_r(\neg p \Rightarrow \neg C)$	$m_r(\top)$
$m_p(p)$	$p \wedge C$	$p \wedge C$	p	p
$m_p(\neg p)$	$\neg p \wedge \neg C$	$\neg p$	$\neg p \wedge \neg C$	$\neg p$
$m_p(\top)$	$p \equiv C$	$p \Rightarrow C$	$\neg p \Rightarrow \neg C$	$\top$

To calculate the belief degree of the conclusion, we sum the masses of all elements that trigger the conclusion (C) noted in green from table 2.2.

$$
\begin{aligned}
Bel_C(C) &= \sum_{\phi:\phi\vdash C,\ \phi\neq\emptyset} m(\phi) = m(p \wedge C) \\
&= m_p(p) \times m_r(p \equiv C) + m_p(p) \times m_r(p \Rightarrow C) \\
&= m_p(p) \cdot [m_r(p \equiv C) + m_r(p \Rightarrow C)] \\
&= m_p(p) \cdot [m_{\Rightarrow}(p \Rightarrow C) \times m_{\Leftarrow}(\neg p \Rightarrow \neg C) + m_{\Rightarrow}(p \Rightarrow C) \times (1 - m_{\Leftarrow}(\neg p \Rightarrow \neg C))] \\
&= m_p(p) \cdot m_{\Rightarrow}(p \Rightarrow C) = Bel_p(p) \cdot Bel_{\Rightarrow}(p \Rightarrow C)
\end{aligned}
$$

To calculate the disbelief degree of the conclusion, we sum the masses of all elements that trigger its negation ($\neg C$) noted in red from table 2.2.

$$
\begin{aligned}
Disb_C(C) &= \sum_{\phi:\phi\vdash \neg C,\ \phi\neq\emptyset} m(\phi) = m(\neg p \wedge \neg C) \\
&= m_p(\neg p) \times m_r(p \equiv C) + m_p(\neg p) \times m_r(\neg p \Rightarrow \neg C) \\
&= m_p(\neg p) \cdot [m_r(p \equiv C) + m_r(\neg p \Rightarrow \neg C)] \\
&= m_p(\neg p) \cdot [m_{\Rightarrow}(p \Rightarrow C) \times m_{\Leftarrow}(\neg p \Rightarrow \neg C) + m_{\Leftarrow}(\neg p \Rightarrow \neg C) \times (1 - m_{\Rightarrow}(p \Rightarrow C))] \\
&= m_p(\neg p) \cdot m_{\Leftarrow}(\neg p \Rightarrow \neg C) = Disb_p(p) \cdot Bel_{\Leftarrow}(\neg p \Rightarrow \neg C)
\end{aligned}
$$

So, we conclude for the uncertainty propagation in simple argument:

$$
\text{S-Arg} : \begin{cases} Bel_C(C) &= Bel_p(p) \cdot Bel_{\Rightarrow}(p \Rightarrow C) \\ Disb_C(C) &= Disb_p(p) \cdot Bel_{\Leftarrow}(\neg p \Rightarrow \neg C) \end{cases} \tag{2.1}
$$

Remarks :

- Since, we work on a two states frame of discernment for premises $\Omega_p = \{p_i, \neg p_i\}$, the conclusion $\Omega_C = \{C, \neg C\}$ and rules, masses and (dis-)belief degrees in this case are equal. For instance, $m_C(C) = Bel_C(C)$ and $m_C(\neg C) = Bel_C(\neg C) = Disb_C(C)$.

- In table 2.2, only the belief in the rules (direct and reverse) was taken into account. Indeed, we decide to not assign a mass on $p \wedge \neg C$ (equivalent to $\neg(p \Rightarrow C)$) and $\neg p \wedge C$ because they are not rules (cannot be expressed with an If-then proposition). Unlike other types of representation (e.g., defeaters in RESSAC graphical notation [Damour 2021]), GSN does not offer the possibility to represent exceptions where one may have a true (resp. false) conclusion despite having a false (resp. true) premise. However, it is allowed to challenge the truth of the premise for valid reasons (e.g, expert cannot trust the experiment because test conditions are not reliable), which is more likely to occur. Therefore, we restrict ourselves only to those rules that can propagate belief and disbelief in the premises.

Notice that in order to calculate the propagation formulas, all masses (on rules and premises) were combined. However, the belief in the conclusion depends only on the direct rule and the belief on the premise (the same remarks are valid on the disbelief in the conclusion). Indeed, reverse rules and the masses on $(\neg p_i)$ cannot infer the acceptance of the conclusion (C) and vise-versa.

Moreover, with complex systems, it is more likely to find goals 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 conclusion. The argument types below describe these situations. All the general formulas will be deduced from the calculation based on a case of a conclusion (C) supported by two premises (p_1) and (p_2).

2.3.2 Conjunctive argument type (C-Arg)

This argument type 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.

For instance, in the example of figure 2.2 both risks (listed in the *Context* box, Ctx) should be treated to guarantee the safety of the battery $(m_{\Rightarrow}([p_1 \wedge p_2] \Rightarrow C), m_{\Rightarrow}(\top))$. However, if at least one premise $(i = 1, 2)$ is not treated, the system is no longer safe $(m^i_{\Leftarrow}(\neg p_i \Rightarrow \neg C), m^i_{\Leftarrow}(\top))$. To get propagation formulas, the mass on rules with those on premises are combined using the conjunctive rule of combination (equation 1.8): $m_C = m_p \otimes m_r$ (tables 2.6 and 2.7) where: $m_p = m_p^1 \otimes m_p^2$ (table 2.3) and $m_r = m_{\Rightarrow} \otimes m^1_{\Leftarrow} \otimes m^2_{\Leftarrow}$ (tables 2.4 and 2.5).

Table 2.3: Combination of the masses on premises

$m_p = m_p^1 \otimes m_p^2$	$m_p^2(p_2)$	$m_p^2(\neg p_2)$	$m_p^2(\top)$
$m_p^1(p_1)$	$p_1 \wedge p_2$	$p_1 \wedge \neg p_2$	p_1
$m_p^1(\neg p_1)$	$\neg p_1 \wedge p_2$	$\neg p_1 \wedge \neg p_2$	$\neg p_1$
$m_p^1(\top)$	p_2	$\neg p_2$	$\top$

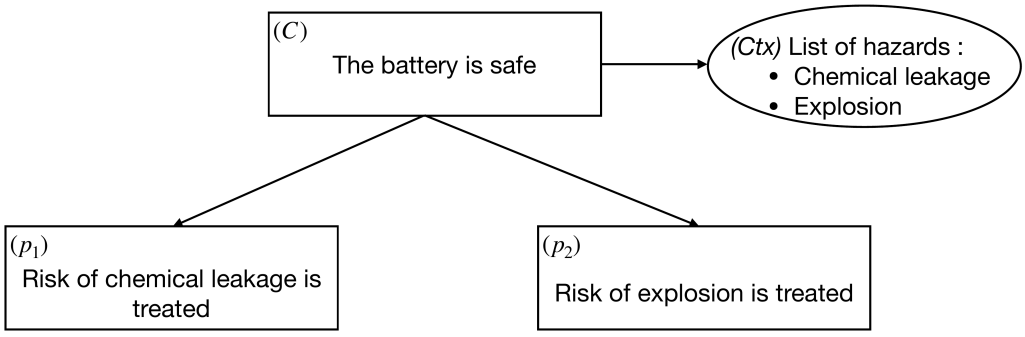


Figure 2.2: A conjunctive argument type (C-Arg) example

Table 2.4: Combination of elementary reverse rules for the C-Arg

$m^{12}_{\Leftarrow} = m^{1}_{\Leftarrow} \otimes m^{2}_{\Leftarrow}$	$m^{2}_{\Leftarrow}(\neg p_2 \Rightarrow \neg C)$	$m^{2}_{\Leftarrow}(\top)$
$m^{1}_{\Leftarrow}(\neg p_1 \Rightarrow \neg C)$	$[\neg p_1 \vee \neg p_2] \Rightarrow \neg C$	$\neg p_1 \Rightarrow \neg C$
$m^{1}_{\Leftarrow}(\top)$	$\neg p_2 \Rightarrow \neg C$	$\top$

Notice that: $[(\neg p_1 \vee \neg p_2) \Rightarrow \neg C] \equiv [\neg(p_1 \wedge p_2) \Rightarrow \neg C]$ (De Morgan's laws)

Table 2.5: Combination of direct and reverse rules for the C-Arg

$m_r = m_{\Rightarrow} \otimes m^{12}_{\Leftarrow}$	$m^{12}_{\Leftarrow}(\neg[p_1 \wedge p_2] \Rightarrow \neg C)$	$m^{12}_{\Leftarrow}(\neg p_1 \Rightarrow \neg C)$	$m^{12}_{\Leftarrow}(\neg p_2 \Rightarrow \neg C)$	$m^{12}_{\Leftarrow}(\top)$
$m_{\Rightarrow}([p_1 \wedge p_2] \Rightarrow C)$	$[p_1 \wedge p_2] \equiv C$	F_1	F_2	$[p_1 \wedge p_2] \Rightarrow C$
$m_{\Rightarrow}(\top)$	$\neg[p_1 \wedge p_2] \Rightarrow \neg C$	$\neg p_1 \Rightarrow \neg C$	$\neg p_2 \Rightarrow \neg C$	$\top$

Where:

- $F_1 = [(p_1 \wedge p_2) \Rightarrow C] \wedge [\neg p_1 \Rightarrow \neg C]$.
- $F_2 = [(p_1 \wedge p_2) \Rightarrow C] \wedge [\neg p_2 \Rightarrow \neg C]$.

Table 2.6: Combination of masses of rules and premises for the C-Arg (part 1)

$m_C = m_p \otimes m_r$	$m_r([p_1 \wedge p_2] \equiv C)$	$m_r([p_1 \wedge p_2] \Rightarrow C)$	$m_r(\neg[p_1 \wedge p_2] \Rightarrow \neg C)$	$m_r(\neg p_1 \Rightarrow \neg C)$
$m_p(p_1 \wedge p_2)$	$p_1 \wedge p_2 \wedge C$	$p_1 \wedge p_2 \wedge C$	-	-
$m_p(\neg p_1 \wedge p_2)$	$\neg p_1 \wedge p_2 \wedge \neg C$	-	$\neg p_1 \wedge p_2 \wedge \neg C$	$\neg p_1 \wedge p_2 \wedge \neg C$
$m_p(p_1 \wedge \neg p_2)$	$p_1 \wedge \neg p_2 \wedge \neg C$	-	$p_1 \wedge \neg p_2 \wedge \neg C$	-
$m_p(\neg p_1 \wedge \neg p_2)$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	-	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$
$m_p(p_1)$	-	-	-	-
$m_p(\neg p_1)$	$\neg p_1 \wedge \neg C$	-	$\neg p_1 \wedge \neg C$	$\neg p_1 \wedge \neg C$
$m_p(p_2)$	-	-	-	-
$m_p(\neg p_2)$	$\neg p_2 \wedge \neg C$	-	$\neg p_2 \wedge \neg C$	-
$m_p(\top)$	-	-	-	-

For a better visualization, we have chosen to replace by (-) the focal elements that do not trigger the conclusion (C) or its negation ($\neg C$) in the tables 2.6 and 2.7.

Table 2.7: Combination of masses of rules and premises for the C-Arg (part 2)

$m_C = m_p \otimes m_r$	$m_r(\neg p_2 \Rightarrow \neg C)$	$m_r(F_1)$	$m_r(F_2)$	$m_r(\top)$
$m_p(p_1 \wedge p_2)$	-	$p_1 \wedge p_2 \wedge C$	$p_1 \wedge p_2 \wedge C$	-
$m_p(\neg p_1 \wedge p_2)$	-	$\neg p_1 \wedge p_2 \wedge \neg C$	-	-
$m_p(p_1 \wedge \neg p_2)$	$p_1 \wedge \neg p_2 \wedge \neg C$	-	$p_1 \wedge \neg p_2 \wedge \neg C$	-
$m_p(\neg p_1 \wedge \neg p_2)$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	-
$m_p(p_1)$	-	-	-	-
$m_p(\neg p_1)$	$\neg p_1 \wedge \neg C$	$\neg p_1 \wedge \neg C$	-	-
$m_p(p_2)$	-	-	-	-
$m_p(\neg p_2)$	$\neg p_2 \wedge \neg C$	-	$\neg p_2 \wedge \neg C$	-
$m_p(\top)$	-	-	-	$\top$

To calculate the belief degree of the conclusion, we sum the masses of all elements that trigger the conclusion (C) noted in green from tables 2.6 and 2.7.

$$\begin{aligned} Bel_C(C) &= \sum_{\phi:\phi\vdash C,\ \phi\neq\emptyset} m(\phi) \\ &= m_C(p_1 \wedge p_2 \wedge C) \\ &= m_p^1(p_1) \cdot m_p^2(p_2) \cdot m_{\Rightarrow}([p_1 \wedge p_2] \Rightarrow C) \\ &= Bel_p^1(p_1) \cdot Bel_p^2(p_2) \cdot Bel_{\Rightarrow}([p_1 \wedge p_2] \Rightarrow C) \end{aligned}$$

To calculate the disbelief degree of the conclusion, we sum the masses of all elements that trigger its negation ($\neg C$) noted in red from tables 2.6 and 2.7.

$$\begin{aligned} Disb_C(C) &= \sum_{\phi:\phi\vdash \neg C,\ \phi\neq\emptyset} m(\phi) \\ &= m_C(\neg p_1 \wedge p_2 \wedge \neg C) + m_C(p_1 \wedge \neg p_2 \wedge \neg C) + m_C(\neg p_1 \wedge \neg p_2 \wedge \neg C) \\ &\quad + m_C(\neg p_1 \wedge \neg C) + m_C(\neg p_2 \wedge \neg C) \\ &= m_p^1(\neg p_1) \cdot m_{\Leftarrow}^1(\neg p_1 \Rightarrow \neg C) + m_p^2(\neg p_2) \cdot m_{\Leftarrow}^2(\neg p_2 \Rightarrow \neg C) \\ &\quad - [m_p^1(\neg p_1) \cdot m_{\Leftarrow}^1(\neg p_1 \Rightarrow \neg C)] \cdot [m_p^2(\neg p_2) \cdot m_{\Leftarrow}^2(\neg p_2 \Rightarrow \neg C)] \\ &= 1 - [1 - m_p^1(\neg p_1) \cdot m_{\Leftarrow}^1(\neg p_1 \Rightarrow \neg C)] \cdot [1 - m_p^2(\neg p_2) \cdot m_{\Leftarrow}^2(\neg p_2 \Rightarrow \neg C)] \\ &= 1 - [1 - Disb_p^1(p_1) \cdot Bel_{\Leftarrow}^1(\neg p_1 \Rightarrow \neg C)] \cdot [1 - Disb_p^2(p_2) \cdot Bel_{\Leftarrow}^2(\neg p_2 \Rightarrow \neg C)] \end{aligned}$$

To compute belief and disbelief degrees of a C-Arg for n premises, we deduce the following formulas (2.2):

$$\text{C-Arg} : \begin{cases} Bel_C(C) &= Bel_{\Rightarrow}([\wedge_{i=1}^n p_i] \Rightarrow C) \cdot \prod_{i=1}^n Bel_p(p_i) \\ Disb_C(C) &= 1 - \prod_{i=1}^n [1 - Disb_p^i(p_i) \cdot Bel_{\Leftarrow}^i(\neg p_i \Rightarrow \neg C)] \end{cases} \tag{2.2}$$

We can notice from the equations in (2.2), that the belief formula takes the form of a general conjunction (the product of belief degrees in premises weighted by the mass on the conjunctive rule). On the other hand, the disbelief formula takes

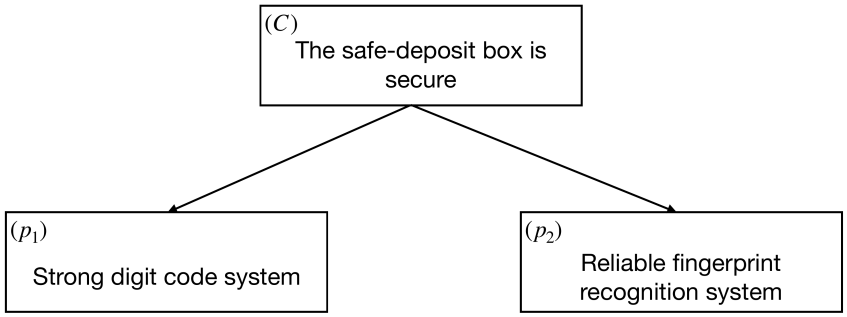


Figure 2.3: A disjunctive argument type (D-Arg) example

the form of a general disjunction. For example, when $n = 2$ premises, we get the probabilistic sum of: $1-(1-Disb^1_C(C))\cdot(1-Disb^2_C(C)) = Disb^1_C(C)+Disb^2_C(C)-Disb^1_C(C)\cdot Disb^2_C(C)$, where: $Disb^i_C(C) = Disb^i_p(p_i)\cdot Bel^i_{\Leftarrow}(\neg p_i \Rightarrow \neg C)$ the disbelief degree in the conclusion induced by the failure of one premise p_i. This argument type favors the propagation of the premise with the least strength (minimal belief due to the product of values between 0 and 1, and maximal disbelief due to the probabilistic sum).

2.3.3 Disjunctive argument type (D-Arg)

This argument type describes the case when the truth of one premise is enough to support the whole conclusion. The corresponding rules are: $\wedge^n_i(p_i \Rightarrow C)$ (direct), and $(\wedge^n_i \neg p_i) \Rightarrow \neg C$ (reverse).

In the example of figure 2.3, each premise (digit code or fingerprint recognition systems) can guarantee the security of the safe-deposit box (direct rules: $m^i_{\Rightarrow}(p_i \Rightarrow C), m^i_{\Rightarrow}(\top)$). Their conjunction does not, in any case, improve the degree of support in the conclusion (C). When both premises p_1 and p_2 are false, then the conclusion is also false, hence the reverse rules: $m_{\Rightarrow}([\neg p_1 \wedge \neg p_2] \Leftarrow \neg C), m_{\Leftarrow}(\top)$. To get propagation formulas, the mass on rules with those on premises are combined using the conjunctive rule of combination (equation 1.8): $m_C = m_p \otimes m_r$ (tables 2.10 and 2.11) where: $m_p = m^1_p \otimes m^2_p$ (table 2.3) and $m_r = m^1_{\Rightarrow} \otimes m^2_{\Rightarrow} \otimes m_{\Leftarrow}$ (tables 2.8 and 2.9).

Table 2.8: Combination of elementary direct rules for the D-Arg

$m^{12}_{\Rightarrow} = m^1_{\Rightarrow} \otimes m^2_{\Rightarrow}$	$m^2_{\Rightarrow}(p_2 \Rightarrow C)$	$m^2_{\Rightarrow}(\top)$
$m^1_{\Rightarrow}(p_1 \Rightarrow C)$	$[p_1 \vee p_2] \Rightarrow C$	$p_1 \Rightarrow C$
$m^1_{\Rightarrow}(\top)$	$p_2 \Rightarrow C$	$\top$

Notice that: $[(\neg p_1 \wedge \neg p_2) \Rightarrow \neg C] \equiv [\neg(p_1 \vee p_2) \Rightarrow \neg C]$ (De Morgan's laws)

Table 2.9: Combination of direct and reverse rules for the D-Arg

$m_r = m_{\Rightarrow} \otimes m^{12}_{\Rightarrow}$	$m^{12}_{\Rightarrow}([p_1 \vee p_2] \Rightarrow C)$	$m^{12}_{\Rightarrow}(p_1 \Rightarrow C)$	$m^{12}_{\Rightarrow}(p_2 \Rightarrow C)$	$m^{12}_{\Rightarrow}(\top)$
$m_{\Leftarrow}(\neg[p_1 \vee p_2] \Rightarrow \neg C)$	$[p_1 \vee p_2] \equiv C$	F'_1	F'_2	$\neg[p_1 \vee p_2] \Rightarrow \neg C$
$m_{\Rightarrow}(\top)$	$[p_1 \vee p_2] \Rightarrow C$	$p_1 \Rightarrow C$	$p_2 \Rightarrow C$	$\top$

Where:

- $F'_1 = [\neg(p_1 \wedge p_2) \Rightarrow \neg C] \wedge [p_1 \Rightarrow C]$.
- $F'_2 = [\neg(p_1 \wedge p_2) \Rightarrow \neg C] \wedge [p_2 \Rightarrow C]$.

Table 2.10: Combination of masses of rules and premises for the D-Arg (part 1)

$m_C = m_p \otimes m_r$	$m_r([p_1 \vee p_2] \equiv C)$	$m_r([p_1 \vee p_2] \Rightarrow C)$	$m_r(\neg[p_1 \vee p_2] \Rightarrow \neg C)$	$m_r(p_1 \Rightarrow C)$
$m_p(p_1 \wedge p_2)$	$p_1 \wedge p_2 \wedge C$	$p_1 \wedge p_2 \wedge C$	-	$p_1 \wedge p_2 \wedge C$
$m_p(\neg p_1 \wedge p_2)$	$\neg p_1 \wedge p_2 \wedge C$	$\neg p_1 \wedge p_2 \wedge C$	-	-
$m_p(p_1 \wedge \neg p_2)$	$p_1 \wedge \neg p_2 \wedge C$	$p_1 \wedge \neg p_2 \wedge C$	-	$p_1 \wedge \neg p_2 \wedge C$
$m_p(\neg p_1 \wedge \neg p_2)$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	-	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	-
$m_p(p_1)$	$p_1 \wedge C$	$p_1 \wedge C$	-	$p_1 \wedge C$
$m_p(\neg p_1)$	-	-	-	-
$m_p(p_2)$	$p_2 \wedge C$	$p_2 \wedge C$	-	
$m_p(\neg p_2)$	-	-	-	-
$m_p(\top)$	-	-	-	-

Table 2.11: Combination of masses of rules and premises for the D-Arg (part 2)

$m_C = m_p \otimes m_r$	$m_r(p_2 \Rightarrow C)$	$m_r(F'_1)$	$m_r(F'_2)$	$m_r(\top)$
$m_p(p_1 \wedge p_2)$	$p_1 \wedge p_2 \wedge C$	$p_1 \wedge p_2 \wedge C$	$p_1 \wedge p_2 \wedge C$	-
$m_p(\neg p_1 \wedge p_2)$	$\neg p_1 \wedge p_2 \wedge C$	-	$\neg p_1 \wedge p_2 \wedge C$	-
$m_p(p_1 \wedge \neg p_2)$	-	$p_1 \wedge \neg p_2 \wedge C$	-	-
$m_p(\neg p_1 \wedge \neg p_2)$	-	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	-
$m_p(p_1)$	-	$p_1 \wedge C$	-	-
$m_p(\neg p_1)$	-	-	-	-
$m_p(p_2)$	$p_2 \wedge C$	-	$p_2 \wedge C$	-
$m_p(\neg p_2)$	-	-	-	-
$m_p(\top)$	-	-	-	$\top$

To calculate the belief degree of the conclusion, we sum the masses of all elements that trigger the conclusion (C) noted in green from tables 2.10 and 2.11.

$$
\begin{aligned}
Bel_C(C) &= \sum_{\phi:\phi\vdash C,\ \phi\neq\emptyset} m(\phi) \\
&= m_C(p_1 \wedge p_2 \wedge C) + m_C(\neg p_1 \wedge p_2 \wedge C) + m_C(p_1 \wedge \neg p_2 \wedge C) + m_C(p_1 \wedge C) + m_C(p_2 \wedge C) \\
&= m^1_p(p_1) \cdot m^1_{\Rightarrow}(p_1 \Rightarrow C) + m^2_p(p_2) \cdot m^2_{\Rightarrow}(p_2 \Rightarrow C) \\
&\quad - [m^1_p(p_1) \cdot m^1_{\Rightarrow}(p_1 \Rightarrow C)] \cdot [m^2_p(p_2) \cdot m^2_{\Rightarrow}(p_2 \Rightarrow C)] \\
&= 1 - [1 - m^1_p(p_1) \cdot m^1_{\Rightarrow}(p_1 \Rightarrow C)] \cdot [1 - m^2_p(p_2) \cdot m^2_{\Rightarrow}(p_2 \Rightarrow C)] \\
&= 1 - [1 - Bel^1_p(p_1) \cdot Bel^1_{\Rightarrow}(p_1 \Rightarrow C)] \cdot [1 - Bel^2_p(p_2) \cdot Bel^2_{\Rightarrow}(p_2 \Rightarrow C)]
\end{aligned}
$$

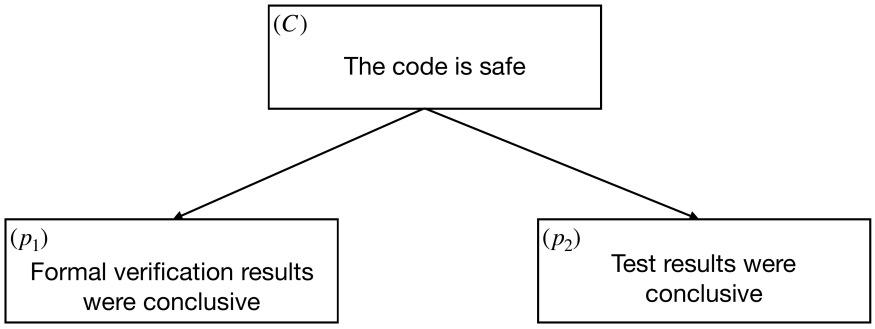


Figure 2.4: A hybrid argument type (H-Arg) example

To calculate the disbelief degree of the conclusion, we sum the masses of all elements that trigger its negation ($\neg C$) noted in red from tables 2.10 and 2.11.

$$\begin{aligned} Disb_C(C) &= \sum_{\phi:\phi\vdash\neg C,\ \phi\neq\emptyset} m(\phi) \\ &= m_C(\neg p_1 \wedge \neg p_2 \wedge \neg C) \\ &= m_p^1(\neg p_1) \cdot m_p^2(\neg p_2) \cdot m_{\Rightarrow}(\neg[p_1 \vee p_2] \Rightarrow \neg C) \\ &= m_p^1(\neg p_1) \cdot m_p^2(\neg p_2) \cdot m_{\Rightarrow}([\neg p_1 \wedge \neg p_2] \Rightarrow \neg C) \\ &= Disb_p^1(p_1) \cdot Disb_p^2(p_2) \cdot Bel_{\Leftarrow}([\neg p_1 \wedge \neg p_2] \Rightarrow \neg C) \end{aligned}$$

To compute belief and disbelief degrees of a D-Arg for n premises, we get the following formulas (2.3):

$$\text{D-Arg} : \begin{cases} Bel_C(C) &= 1 - \prod_{i=1}^n [1 - Bel_p^i(p_i) \cdot Bel_{\Rightarrow}^i(p_i \Rightarrow C)] \\ Disb_C(C) &= Bel_{\Leftarrow}([\wedge_{i=1}^n \neg p_i] \Rightarrow \neg C) \cdot \prod_{i=1}^n Disb_p^i(p_i) \end{cases} \tag{2.3}$$

In opposition to the C-Arg, the belief (resp. disbelief) formula in (2.3) expresses a general disjunction (resp. conjunction). This argument (D-Arg) favors the propagation of the premise with the greatest strength (maximal belief and minimal disbelief).

2.3.4 Hybrid argument type (H-Arg)

This argument type describes the case where each premise supports the conclusion to some extent $\wedge_i^n(p_i \Rightarrow C)$, but their conjunction does it to a larger extent $(\wedge_i^n p_i) \Rightarrow C$ (direct rules). 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. In the example of figure 2.4, the premise "*Test results were conclusive*" supports the conclusion to some point. Since, evidence on formal verification was also provided, which allows to identify some unsafe states that the system will never reach, experts usually conduct limited tests (which are limited by issues such as cost, feasibility, etc). On the other hand, tests can cover issues that

formal verification might not capture. Unlike the D-Arg, the conjunction of these two premises improves the support degree in the conclusion. To get propagation formulas in 2.4, we combine the mass on rules (elementary and conjunctive, direct and reverse) with those on premises using the conjunctive rule of combination, i.e: $m_C = m_p \otimes m_r$ (tables 2.16, 2.16 and 2.17) where: $m_p = m_p^1 \otimes m_p^2$ (table 2.3) and $m_r = m_{\Leftarrow} \otimes m_{\Leftarrow}^1 \otimes m_{\Leftarrow}^2 \otimes m_{\Rightarrow} \otimes m_{\Rightarrow}^1 \otimes m_{\Rightarrow}^2$ (tables 2.12, 2.13, 2.14).

Table 2.12: Combination of direct and reverse conjunctive rules for H-Arg

$m_r' = m_{\Rightarrow} \otimes m_{\Leftarrow}$	$m_{\Rightarrow}([\neg p_1 \wedge \neg p_2] \Rightarrow \neg C)$	$m_{\Rightarrow}(\top)$
$m_{\Rightarrow}([p_1 \wedge p_2] \Rightarrow C)$	$F_c = ([p_1 \wedge p_2] \Rightarrow C) \wedge ([\neg p_1 \wedge \neg p_2] \Rightarrow \neg C)$	$[p_1 \wedge p_2] \Rightarrow C$
$m_{\Rightarrow}(\top)$	$[\neg p_1 \wedge \neg p_2] \Rightarrow \neg C$	$\top$

Table 2.13: Combination of direct and reverse elementary rules for the H-Arg

$m_r^i = m_{\Rightarrow}^i \otimes m_{\Leftarrow}^i$	$m_{\Rightarrow}^i(p_i \Rightarrow C)$	$m_{\Rightarrow}^i(\top)$
$m_{\Leftarrow}^i(\neg p_i \Rightarrow \neg C)$	$p_i \equiv C$	$\neg p_i \Rightarrow \neg C$
$m_{\Leftarrow}^i(\top)$	$p_i \Rightarrow C$	$\top$

Table 2.14: Combination of premise (p_i) with its elementary rules for H-Arg

$m_i = m_p^i \otimes m_r^i$	$m_r^i(p_i \equiv C)$	$m_r^i(p_i \Rightarrow C)$	$m_r^i(\neg p_i \Rightarrow \neg C)$	$m_r^i(\top)$
$m_p^i(p_i)$	$p_i \wedge C$	$p_i \wedge C$	p_i	p_i
$m_p^i(\neg p_i)$	$\neg p_i \wedge \neg C$	$\neg p_i$	$\neg p_i \wedge \neg C$	$\neg p_i$
$m_p^i(\top)$	$p_i \equiv C$	$p_i \Rightarrow C$	$\neg p_i \Rightarrow \neg C$	$\top$

Table 2.15: Combination of premise (p_i) with its elementary rules of two premises (1 and 2) for H-Arg (part 1)

$m_{12} = m_1 \otimes m_2$	$m_2(p_2 \wedge C)$	$m_2(\neg p_2 \wedge \neg C)$	$m_2(p_2)$	$m_2(\neg p_2)$
$m_1(p_1 \wedge C)$	$p_1 \wedge p_2 \wedge C$	$\emptyset$	$p_1 \wedge p_2 \wedge C$	$p_1 \wedge \neg p_2 \wedge C$
$m_1(\neg p_1 \wedge \neg C)$	$\emptyset$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	$\neg p_1 \wedge p_2 \wedge \neg C$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$
$m_1(p_1)$	$p_1 \wedge p_2 \wedge C$	$p_1 \wedge \neg p_2 \wedge \neg C$	$p_1 \wedge p_2$	$p_1 \wedge \neg p_2$
$m_1(\neg p_1)$	$\neg p_1 \wedge p_2 \wedge C$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	$\neg p_1 \wedge p_2$	$\neg p_1 \wedge \neg p_2$
$m_1(p_1 \Rightarrow C)$	$p_2 \wedge C$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	$p_2(p_1 \Rightarrow C)$	$\neg p_2(p_1 \Rightarrow C)$
$m_1(\neg p_1 \Rightarrow \neg C)$	$p_1 \wedge p_2 \wedge C$	$\neg p_2 \wedge \neg C$	$p_2(\neg p_1 \Rightarrow \neg C)$	$\neg p_2(\neg p_1 \Rightarrow \neg C)$
$m_1(p_1 \equiv C)$	$p_1 \wedge p_2 \wedge C$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	$p_2(p_1 \equiv C)$	$\neg p_2(p_1 \equiv C)$
$m_1(\top)$	$p_1 \wedge C$	$\neg p_2 \wedge \neg C$	p_2	$\neg p_2$

Table 2.16: Combination of premise (p_i) with its elementary rules of two premises (1 and 2) for H-Arg (part 2)

$m_{12} = m_1 \otimes m_2$	$m_2(p_2 \Rightarrow C)$	$m_2(\neg p_2 \Rightarrow \neg C)$	$m_2(p_2 \equiv C)$	$m_2(\top)$
$m_1(p_1 \wedge C)$	$p_1 \wedge p_2 \wedge C$	$p_1 \wedge p_2 \wedge C$	$p_1 \wedge C$	$p_1 \wedge C$
$m_1(\neg p_1 \wedge \neg C)$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	$\neg p_1 \wedge \neg C$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	$\neg p_1 \wedge \neg C$
$m_1(p_1)$	$p_1(p_2 \Rightarrow C)$	$p_1(\neg p_2 \Rightarrow \neg C)$	$p_1(p_2 \equiv C)$	p_1
$m_1(\neg p_1)$	$\neg p_1(p_2 \Rightarrow C)$	$\neg p_1(\neg p_2 \Rightarrow \neg C)$	$\neg p_1(p_2 \equiv C)$	$\neg p_1$
$m_1(p_1 \Rightarrow C)$	$(p_1 \Rightarrow C)(p_2 \Rightarrow C)$	$(p_1 \Rightarrow C)(\neg p_2 \Rightarrow \neg C)$	$(p_1 \Rightarrow C)(p_2 \equiv C)$	$(p_1 \Rightarrow C)$
$m_1(\neg p_1 \Rightarrow \neg C)$	$(p_2 \Rightarrow C)(\neg p_1 \Rightarrow \neg C)$	$(\neg p_2 \Rightarrow \neg C)(\neg p_1 \Rightarrow \neg C)$	$(p_2 \equiv C)(\neg p_1 \Rightarrow \neg C)$	$\neg p_1 \Rightarrow \neg C$
$m_1(p_1 \equiv C)$	$(p_1 \equiv C)(p_2 \Rightarrow C)$	$(p_1 \equiv C)(\neg p_2 \Rightarrow \neg C)$	$(p_1 \equiv C)(p_2 \equiv C)$	$(p_1 \equiv C)$
$m_1(\top)$	$p_2 \Rightarrow C$	$\neg p_2 \Rightarrow \neg C$	$p_2 \equiv C$	$\top$

Table 2.17: Combination of premise (p_i) with its rules for H-Arg

$m_C = m_{12} \otimes m'_r$	$m_r(F_c)$	$m_r([p_1 \wedge p_2] \Rightarrow C)$	$m_r([\neg p_1 \wedge \neg p_2] \Rightarrow \neg C)$	$m_r(\top)$
$m_{12}(p_1 \wedge p_2 \wedge C)$	$p_1 \wedge p_2 \wedge C$	$p_1 \wedge p_2 \wedge C$	$p_1 \wedge p_2 \wedge C$	$p_1 \wedge p_2 \wedge C$
$m_{12}(\neg p_1 \wedge p_2 \wedge C)$	$\neg p_1 \wedge p_2 \wedge C$	$\neg p_1 \wedge p_2 \wedge C$	$\neg p_1 \wedge p_2 \wedge C$	$\neg p_1 \wedge p_2 \wedge C$
$m_{12}(p_1 \wedge \neg p_2 \wedge C)$	$p_1 \wedge \neg p_2 \wedge C$	$p_1 \wedge \neg p_2 \wedge C$	$p_1 \wedge \neg p_2 \wedge C$	$p_1 \wedge \neg p_2 \wedge C$
$m_{12}(p_1 \wedge C)$	$p_1 \wedge C$	$p_1 \wedge C$	$p_1 \wedge C$	$p_1 \wedge C$
$m_{12}(p_2 \wedge C)$	$p_2 \wedge C$	$p_2 \wedge C$	$p_2 \wedge C$	$p_2 \wedge C$
$m_{12}(\neg p_1 \wedge \neg p_2 \wedge \neg C)$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$
$m_{12}(\neg p_1 \wedge p_2 \wedge \neg C)$	$\neg p_1 \wedge p_2 \wedge \neg C$	$\neg p_1 \wedge p_2 \wedge \neg C$	$\neg p_1 \wedge p_2 \wedge \neg C$	$\neg p_1 \wedge p_2 \wedge \neg C$
$m_{12}(p_1 \wedge \neg p_2 \wedge \neg C)$	$p_1 \wedge \neg p_2 \wedge \neg C$	$p_1 \wedge \neg p_2 \wedge \neg C$	$p_1 \wedge \neg p_2 \wedge \neg C$	$p_1 \wedge \neg p_2 \wedge \neg C$
$m_{12}(\neg p_1 \wedge \neg C)$	$\neg p_1 \wedge \neg C$	$\neg p_1 \wedge \neg C$	$\neg p_1 \wedge \neg C$	$\neg p_1 \wedge \neg C$
$m_{12}(\neg p_2 \wedge \neg C)$	$\neg p_2 \wedge \neg C$	$\neg p_2 \wedge \neg C$	$\neg p_2 \wedge \neg C$	$\neg p_2 \wedge \neg C$
$m_{12}(p_1 \wedge p_2)$	$p_1 \wedge p_2 \wedge C$	$p_1 \wedge p_2 \wedge C$	-	-
$m_{12}(\neg p_1 \wedge \neg p_2)$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	-	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	-
...	...	...	...	...
$m_{12}(\emptyset)$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
$m_{12}(\top)$	F_c	$[p_1 \wedge p_2] \Rightarrow C$	$[\neg p_1 \wedge \neg p_2] \Rightarrow \neg C$	$\top$

In table 2.17, we have chosen not to represent all the focal elements resulting from the tables 2.15 and 2.16 which does not trigger the conclusion or its negation to reduce the size of the table.

To calculate the belief degree of the conclusion, we sum the masses of all elements that trigger the conclusion (C) noted in green from table 2.17:

$$\begin{aligned}
Bel_C(C) &= m_C(p_1 \wedge p_2 \wedge C) + m_C(\neg p_1 \wedge p_2 \wedge C) + m_C(p_1 \wedge \neg p_2 \wedge C) + m_C(p_1 \wedge C) + m_C(p_2 \wedge C) - m_C(\emptyset) \\
&= \{m_p^1(p_1) \cdot m_\Rightarrow^1(p_1 \Rightarrow C) + m_p^2(p_2) \cdot m_\Rightarrow^2(p_2 \Rightarrow C) - [m_p^1(p_1) \cdot m_\Rightarrow^1(p_1 \Rightarrow C)] \cdot [m_p^2(p_2) \cdot m_\Rightarrow^2(p_2 \Rightarrow C)]\} \\
&\quad + \{m_p^1(p_1) \cdot m_p^2(p_2) \cdot m_\Rightarrow([p_1 \wedge p_2] \Rightarrow C) \cdot [1 - m_\Rightarrow^1(p_1 \Rightarrow C)] \cdot [1 - m_\Rightarrow^2(p_2 \Rightarrow C)]\} - m_C(\emptyset) \\
&= \{1 - [1 - m_p^1(p_1) \cdot m_\Rightarrow^1(p_1 \Rightarrow C)] \cdot [1 - m_p^2(p_2) \cdot m_\Rightarrow^2(p_2 \Rightarrow C)]\} \\
&\quad + \{m_p^1(p_1) \cdot m_p^2(p_2) \cdot m_\Rightarrow([p_1 \wedge p_2] \Rightarrow C) \cdot [1 - m_\Rightarrow^1(p_1 \Rightarrow C)] \cdot [1 - m_\Rightarrow^2(p_2 \Rightarrow C)]\} - m_C(\emptyset) \\
&= \{1 - [1 - Bel_p^1(p_1) \cdot Bel_\Rightarrow^1(p_1 \Rightarrow C)] \cdot [1 - Bel_p^2(p_2) \cdot Bel_\Rightarrow^2(p_2 \Rightarrow C)]\} \\
&\quad + \{Bel_p^1(p_1) \cdot Bel_p^2(p_2) \cdot Bel_\Rightarrow([p_1 \wedge p_2] \Rightarrow C) \cdot [1 - Bel_\Rightarrow^1(p_1 \Rightarrow C)] \cdot [1 - Bel_\Rightarrow^2(p_2 \Rightarrow C)]\} - m_C(\emptyset)
\end{aligned}$$

To calculate the disbelief degree of the conclusion, we sum the masses of all elements that trigger its negation ($\neg C$) noted in red from table 2.17:

$$\begin{aligned}
Disb_C(C) &= m_C(\neg p_1 \wedge \neg p_2 \wedge \neg C) + m_C(\neg p_1 \wedge p_2 \wedge \neg C) + m_C(p_1 \wedge \neg p_2 \wedge \neg C) + m_C(\neg p_1 \wedge \neg C) + m_C(\neg p_2 \wedge \neg C) - m(\emptyset) \\
&= \{m_p^1(\neg p_1) \cdot m_\Leftarrow^1(\neg p_1 \Rightarrow \neg C) + m_p^2(\neg p_2) \cdot m_\Leftarrow^2(\neg p_2 \Rightarrow \neg C) - [m_p^1(\neg p_1) \cdot m_\Leftarrow^1(\neg p_1 \Rightarrow \neg C)] \cdot [m_p^2(\neg p_2) \cdot m_\Leftarrow^2(\neg p_2 \Rightarrow \neg C)]\} \\
&\quad + \{m_p^1(\neg p_1) \cdot m_p^2(\neg p_2) \cdot m_\Leftarrow([\neg p_1 \wedge \neg p_2] \Rightarrow \neg C) \cdot [1 - m_\Leftarrow^1(\neg p_1 \Rightarrow \neg C)] \cdot [1 - m_\Leftarrow^2(\neg p_2 \Rightarrow \neg C)]\} - m_C(\emptyset) \\
&= \{1 - [1 - m_p^1(\neg p_1) \cdot m_\Leftarrow^1(\neg p_1 \Rightarrow \neg C)] \cdot [1 - m_p^2(\neg p_2) \cdot m_\Leftarrow^2(\neg p_2 \Rightarrow \neg C)]\} \\
&\quad + \{m_p^1(\neg p_1) \cdot m_p^2(\neg p_2) \cdot m_\Leftarrow([\neg p_1 \wedge \neg p_2] \Rightarrow \neg C) \cdot [1 - m_\Leftarrow^1(\neg p_1 \Rightarrow \neg C)] \cdot [1 - m_\Leftarrow^2(\neg p_2 \Rightarrow \neg C)]\} - m_C(\emptyset) \\
&= \{1 - [1 - Disb_p^1(p_1) \cdot Bel_\Leftarrow^1(\neg p_1 \Rightarrow \neg C)] \cdot [1 - Disb_p^2(p_2) \cdot Bel_\Leftarrow^2(\neg p_2 \Rightarrow \neg C)]\} \\
&\quad + \{Disb_p^1(p_1) \cdot Disb_p^2(p_2) \cdot Bel_\Leftarrow([\neg p_1 \wedge \neg p_2] \Leftarrow \neg C) \cdot [1 - Bel_\Leftarrow^1(\neg p_1 \Leftarrow \neg C)] \cdot [1 - Bel_\Leftarrow^2(\neg p_2 \Rightarrow \neg C)]\} - m_C(\emptyset)
\end{aligned}$$

To calculate the mass of conflict for two premises, we sum the masses of all elements that trigger the empty set ($\emptyset$) noted in blue from table 2.17:

$$
\begin{aligned}
m_C(\emptyset) &= m_{12}(\emptyset) \\
&= m_1(\neg p_1 \wedge \neg C) \cdot m_2(p_2 \wedge \neg C) + m_2(\neg p_2 \wedge \neg C) \cdot m_1(p_1 \wedge \neg C) \\
&= m_p^1(\neg p_1) \cdot m_{\Leftarrow}^1(\neg p_1 \Rightarrow \neg C) \cdot Bel_p^2(p_2) \cdot m_{\Rightarrow}^2(p_2 \Rightarrow C) + m_p^2(\neg p_2) \cdot m_{\Leftarrow}^2(\neg p_2 \Rightarrow \neg C) \cdot m_p^1(p_1) \cdot Bel_{\Rightarrow}^1(p_1 \Rightarrow C) \\
&= Disb_p^1(p_1) \cdot Bel_{\Leftarrow}^1(\neg p_1 \Rightarrow \neg C) \cdot Bel_p^2(p_2) \cdot Bel_{\Rightarrow}^2(p_2 \Rightarrow C) + Disb_p^2(p_2) \cdot Bel_{\Leftarrow}^2(\neg p_2 \Rightarrow \neg C) \cdot Bel_p^1(p_1) \cdot Bel_{\Rightarrow}^1(p_1 \Rightarrow C)
\end{aligned}
$$

To compute belief and disbelief degrees of a H-Arg for n premises, we get the following formulas (2.4):

$$\text{H-Arg} : \begin{cases} Bel_C(C) & = Bel_{\Rightarrow}([\wedge_{i=1}^{n} p_i] \Rightarrow C) \times \prod_{i=1}^{n} Bel_p^i(p_i) \cdot [1 - Bel_{\Rightarrow}^i(p_i \Rightarrow C)] \\ & +\{1 - \prod_{i=1}^{n}[1 - Bel_p^i(p_i) \cdot Bel_{\Rightarrow}^i(p_i \Rightarrow C)]\} - m_C^{(n)}(\perp) \\ Disb_C(C) & = Bel_{\Leftarrow}([\wedge_{i=1}^{n} \neg p_i] \Rightarrow \neg C) \times \prod_{i=1}^{n} Disb_p^i(p_i) \cdot [1 - Bel_{\Leftarrow}^i(\neg p_i \Rightarrow \neg C)] \\ & +\{1 - \prod_{i=1}^{n}[1 - Disb_p^i(p_i) \cdot Bel_{\Leftarrow}^i(\neg p_i \Rightarrow \neg C)]\} - m_C^{(n)}(\perp) \end{cases} \tag{2.4}$$

Where the mass $m_C^{(n)}(\perp)$ represents the conflict degree for (n) premises. This mass will be defined in the next section

We can notice that the formulas in (2.4), have the same form. They sum two parts. A first one which expresses the conjunctive component of this type, weighted by the uncertainty on direct elementary rules $(1 - Bel_{\Rightarrow}^i(p_i \Rightarrow C)$, resp. the reverse one). The second part represent its disjunctive component. To deduce the formulas of a C-Arg from H-Arg, all you need is to set the elementary rules $(Bel_{\Rightarrow}^i(p_i \Rightarrow C))$ and the reverse conjunctive one $(Bel_{\Leftarrow}([\wedge_{i=1}^{n} \neg p_i] \Rightarrow \neg C))$ to zero. Conversely, if we set to zero the masses of the direct conjunctive rule and those of the reverse elementary rules, we get the formulas of a D-Arg. Finally, in the case of a conclusion supported by one premise both masses of the direct and reverse conjunctive rules are equal to zero. Thus, we also deduce the formula of an S-Arg.

Remark : The hybrid type presented in this paper, also represents a special case of a situation where one or more premises can support the conclusion to a certain degree, and their conjunction does it to a higher degree. For example, in the case of a conclusion (C) supported by three premises (p_1, p_2 and p_3) we will have seven direct rules ($\{p_1$; p_2; p_3; $p_1 \wedge p_2$; $p_1 \wedge p_3$; $p_2 \wedge p_3$ and $p_1 \wedge p_2 \wedge p_3\}$ implies C) and seven more for the reverse rules. The formulas of such an argument can be calculated following the same method as above. Despite the computational difficulty (more focal sets to combine), such a calculation is feasible. However, the resulting model will require more information (belief degrees) which are difficult to obtain. This issue will be put into perspective for future work, where it will be interesting to know if this trade-off between rigorous modeling of argument types and the complexity of collecting data was worth it.

2.3.5 Conflict mass for (n) premises

A conflict situation represents the case when one or more premises lead to opposite assessments of the conclusion (e.g., a premise p_i supports a conclusion C, while another premise p_j supports its negation). Formally, it always takes the form of a combination of four items : p_i, $p_i \Rightarrow C$, $\neg p_j$ and $\neg p_j \Rightarrow \neg C$, which trigger empty intersections (noticeable if we combine the masses of elementary rules and those on the premises, e.g. 2.15 and 2.16). In the case of the C-Arg and D-Arg this combination never occurs. Indeed, since the definition of a

C-Arg (resp. D-Arg) does not define direct (resp. reverse) elementary rules that combine with the reverse (resp. direct) elementary rules will trigger an empty intersection. Thus we can say that the value of $m_C^{(n)}(\perp)$ is always zero for these types.

Example: Let us consider two premises p_i and p_j with opposite assessments $(Bel_p^i(p_i) = 0,\ Disb_p^i(p_i) = 1)$ and $(Bel_p^j(p_j) = 1,\ Disb_p^j(p_j) = 0)$.

1. In the case of a C-Arg for n premises with maximal beliefs of rules, we obtain:

$$\begin{cases} Bel_C(C) & = Bel_p^1(p_1) \times \cdots \times Bel_p^i(p_i) \times \cdots \times Bel_p^n(p_n) = 0, \quad Bel_p^i(p_i) = 0 \\ Disb_C(C) & = 1 - [1 - Disb_p^1(p_1)] \times \cdots \times [1 - Disb_p^j(p_j)] \times \cdots \times [1 - Disb_p^n(p_n)] \\ & = 1 - 0 = 1 \end{cases}$$

2. In the case of a D-Arg for n premises with maximal beliefs of rules, we obtain:

$$\begin{cases} Bel_C(C) & = 1 - \prod_{k=1}^{n}[1 - Bel_p^k(p_k)] = 1, \quad \text{when } k = i,\ Bel_p^j(p_j) = 1 \\ Disb_C(C) & = \prod_{k=1}^{n} Disb_p^k(p_k) = 0, \quad \text{when } k = j,\ Disb_p^j(p_j) = 0 \end{cases}$$

Facing a conflict situation, when we have opposite assessment on premises supporting the same goal, the conjunctive type (1) adopts a cautious behavior in favor of the propagation of the premises that does not support the conclusion. On the contrary, the disjunctive type (2) takes a more optimistic view, which favors the propagation of the premises that support the conclusion.

Standing between these two limit cases, the value of $m_C^{(n)}(\perp)$[2] for the hybrid type (H-Arg) may have a positive value.

We propose in the following a recursive formula to calculate conflict mass for $n \geq 2$ premises:

$$m_C^{(n)}(\perp) = Bel_C^{(n-1)}(C) \times m_n(\neg p_n \wedge \neg C) + Disb_C^{(n-1)}(C) \times m_n(p_n \wedge C) + m_C^{(n-1)}(\perp) \tag{2.5}$$

Where :

$m_C^{(1)}(\perp) = 0$

$Bel_C^{(n-1)}(C) = \{1 - \prod_{i=1}^{n-1}[1 - Bel_p^i(p_i) \cdot Bel_{\Rightarrow}^i(p_i \Rightarrow C)]\} - m_C^{(n-1)}(\emptyset)$

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

$m_i(p_i \wedge C) = Bel_p^i(p_i) \cdot Bel_{\Rightarrow}^i(p_i \Rightarrow C)$

$m_i(\neg p_i \wedge \neg C) = Disb_p^i(p_i) \cdot Bel_{\Leftarrow}^i(\neg p_i \Rightarrow \neg C)$

Proof : Let us enumerate all conflicting combinations of 4 formulas when merging mass functions on premises and rules.

For $n = 2$, we get two contradictory combinations $p_1 \wedge (p_1 \Rightarrow C) \wedge \neg p_2 \wedge (\neg p_2 \Rightarrow \neg C)$ and $\neg p_1 \wedge (\neg p_1 \Rightarrow \neg C) \wedge p_2 \wedge (p_2 \Rightarrow C)$. So, we calculate the mass on the

[2] Also, noted $m(\emptyset)$ according to the set theory syntax : $[\perp] = \emptyset$.

conflict:

$$\begin{aligned} m_{12}(\bot) &= m_1(p_1 \wedge C) \times m_2(\neg p_2 \wedge \neg C) + m_1(\neg p_1 \wedge \neg C) \times m_2(p_2 \wedge C) \\ &= Bel^1_p(p_1) \cdot Bel^1_{\Rightarrow}(p_1 \Rightarrow C) \times m_2(\neg p_2 \wedge \neg C) \\ &\quad + Disb^1_p(p_1) \cdot Bel^1_{\Leftarrow}(\neg p_1 \Rightarrow \neg C) \times m_2(p_2 \wedge C) \\ &= Bel_C(C) \times m_2(\neg p_2 \wedge \neg C) + Disb_C(C) \times m_2(p_2 \wedge C) \end{aligned}$$

For $n = 3$, we can calculate $m_{123} = m_1 \otimes m_2 \otimes m_3$ directly but we have to calculate $8 \times 8 \times 8$ intersections and select the contradictory conjunctions. We can also calculate more easily:

$$\begin{aligned} m_{123}(\bot) &= \sum_{i \neq j; i,j=1,2,3} m_i(p_i \wedge C) \cdot m_j(\neg p_j \wedge \neg C) \\ &= m_{12}(\bot) + m_{13}(\bot) + m_{23}(\bot) \end{aligned}$$

However, this calculation counts contradictory terms several times. Thus, we may find that this mass ($m_{123}(\bot)$) is greater than 1. Note the contradictory term $m_1(p_1 \wedge C) \cdot m_2(\neg p_2 \wedge \neg C) \cdot m_3(\neg p_3 \wedge \neg C)$ with $1\bar{2}\bar{3}$, etc. We have $m_1(p_1 \wedge C) \cdot m_2(\neg p_2 \wedge \neg C) \cdot \sum_{\phi \in \Omega_3} m_3(\phi)$, with $\sum_{\phi \in \Omega_3} m_3(\phi) = 1$. Hence, the term $m_1(p_1 \wedge C) \cdot m_2(\neg p_2 \wedge \neg C)$ of $m_{12}(\bot)$ includes $1\bar{2}\bar{3}$ and $1\bar{2}3$. Similarly, we get the remaining duplicate terms (in bold):

- $m_1(\neg p_1 \wedge \neg C) \cdot m_2(p_2 \wedge C)$ includes $\bar{1}2\bar{3}$ et $\bar{1}23$.
- $m_1(\neg p_1 \wedge \neg C) \cdot m_3(p_3 \wedge C)$ includes $\mathbf{\bar{1}23}$ et $\bar{1}\bar{2}3$.
- $m_1(p_1 \wedge C) \cdot m_3(\neg p_3 \wedge \neg C)$ includes $12\bar{3}$ et $\mathbf{1\bar{2}\bar{3}}$.
- $m_2(p_2 \wedge C) \cdot m_3(\neg p_3 \wedge \neg C)$ includes $\mathbf{12\bar{3}}$ et $\mathbf{\bar{1}2\bar{3}}$.
- $m_2(\neg p_2 \wedge \neg C) \cdot m_3(p_3 \wedge C)$ includes $\mathbf{1\bar{2}3}$ et $\mathbf{\bar{1}\bar{2}3}$.

hence, the sum $m_{12}(\bot) + m_{13}(\bot) + m_{23}(\bot)$ counts twice the product of three terms. There are 12 such terms, so we have to delete 6 of them (the ones in bold). Thus we prove that:

$$\begin{aligned} m_{123}(\bot) &= m_{12}(\bot) + m_1(p_1 \wedge C) m_3(\neg p_3 \wedge \neg C)[1 - m_2(\neg p_2 \wedge \neg C)] \\ &\quad + m_1(\neg p_1 \wedge \neg C) m_3(p_3 \wedge C)[1 - m_2(p_2 \wedge C)] \\ &\quad + m_2(p_2 \wedge C) m_3(\neg p_3 \wedge \neg C)[1 - m_1(p_1 \wedge C) - m_1(\neg p_1 \wedge \neg C)] \\ &\quad + m_2(\neg p_2 \wedge \neg C) m_3(p_3 \wedge C)[1 - m_1(p_1 \wedge C) - m_1(\neg p_1 \wedge \neg C)] \\ &= m_{12}(\bot) + m_3(p_3 \wedge C) \cdot Disb^{(2)}_C(C) + m_3(\neg p_3 \wedge \neg C) \cdot Bel^{(2)}_C(C) \end{aligned}$$

Where:

$Bel^{(2)}_C(C) = 1 - [1 - Bel^1_p(p_1) Bel^1_{\Rightarrow}(p_1 \Rightarrow C)] \cdot [1 - Bel^2_p(p_2) Bel^2_{\Rightarrow}(p_2 \Rightarrow C)]$

$Disb^{(2)}_C(C) = 1 - [1 - Disb^1_p(p_1) Bel^1_{\Leftarrow}(\neg p_1 \Rightarrow \neg C)] \cdot [1 - Disb^2_p(p_2) Bel^2_{\Leftarrow}(\neg p_2 \Rightarrow \neg C)]$

This calculation can be extended to $n > 1$ premises. Hence, we get $\sum_{i=1}^{n-1} C_{n-1}^{i}(2^i - 1)$ focal sets inducing the conclusion C ($\{\Omega_{p_1} \times ... \times \Omega_{p_n}\} \wedge C$) and others of the same count inducing its negation $\neg C$ ($\{\Omega_{p_1} \times ... \times \Omega_{p_n}\} \wedge \neg C$). Combining these focal sets respectively with ($\neg p_n \wedge \neg C$) and ($p_n \wedge C$) trigger an empty intersection. Summing the masses of these focal sets gives the general formula of conflict $m^{(n)}(\perp)$ (equation 2.5).

Notice that conjunctive rules $[\wedge_{i=1}^{n} p_i] \Rightarrow C$ and $[\wedge_{i=1}^{n} \neg p_i] \Rightarrow \neg C$ were not involved in this calculation because they don't trigger empty intersection (e.g., $([p_1 \wedge p_2] \Rightarrow C) \otimes \neg p_1 = \neg p_1$).

In the case of a minor conflict ($m(\perp)$ small) normalising by the conflict degree $(1-m(\perp))$ as proposed in the usual DS rule of combination will eliminate the conflict and proportionally increase the contradiction-free degrees of the belief $Bel_C(C)$ and disbelief $Disb_C(C)$ in a misleading way. While in a case of full conflict, this mass is equal to 1 making this normalisation meaningless (division by zero). Keeping $m(\perp)$ provides additional information in the sense that it may show that the system is not so safe because of the presence of a conflict (proof for Vs proof against).

Since the H-Arg represent a general form for the S-Arg, C-Arg and D-Arg, in the following we are going to use only the propagation formulas of the hybrid argument type to calculate all belief and disbelief degrees in the conclusion.

2.4 Argument types in literature - A comparison

As seen in chapter 1, most argument types proposed in literature (e.g., [Cyra 2011, Wang 2018a]) are built around three concepts: conjunction (logical AND), disjunction (logical OR) and a mix between the two. However, the difference between these works, which lead to different uncertainty propagation formulas, resides in several aspects.

Formal definition of arguments

The first is the formal definition of each argument type. Several authors (e.g., [Cyra 2011, Ayoub 2013]) simply give informal definitions expressed in natural language. This type of definition gives indeed a general picture of the propagation schema in those argument types, but it is not sufficient to build accurate propagation models. In Wang et al. works [Wang 2018b, Wang 2019], logical conjunction and disjunction are used to express the relation between premises and logical equivalence ($\equiv$) is used to express the relation between premises and the conclusion. Compared to our work, we choose to replace this equivalence with an implication ($\Rightarrow$). This choice is justified by the fact that equivalence provides information on both belief and disbelief, but the provider of such information (e.g., an expert of the filed) may not know one of the two which cannot be separated. Moreover, an agent (or a client) can be only interested in one side of the assessment (generally the belief side to validate a system, or in some cases the disbelief side for an internal

evaluation of a system). Implication offers the possibility to use both sides (belief and disbelief) separately or together.

Combination rule and mass assignment

The second difference is the choice of a combination rule and mass assignment. Most works we encountered, used weighted average [Ayoub 2013] and Dempster rule of combination to merge the masses on rules and premises. We choose to use only the conjunctive rule of combination because it is simple to apply and consistent with the definitions of our argument types (a conjunction of rules). For instance, a conjunction of $[\wedge_{i=1}^{n} p_i] \Rightarrow C$ and $\wedge_{i=1}^{n}[\neg p_i \Rightarrow \neg C]$ for the C-Arg. On the other hand, we choose to assign one mass function per rule and premise because otherwise they become dependent on each other ($\sum_{E \subseteq \Omega} m(E) = 1$). In Wang et al. works the mass of the conjunction ($[\wedge_{i=1}^{n} p_i] \equiv C$) and disjunction ($[\vee_{i=1}^{n} p_i] \equiv C$) is calculated based on elementary rules ($\wedge_{i=1}^{n}[p_i \equiv C]$). If the mass of one rule is maximal, the masses on the other rules are automatically minimal. Therefore, the effect of all the premises on the conclusion (rules) is neglected in favour of the one with the highest weight. Moreover, this configuration (one mass function for all rules) does not reveal the conflict that may exist in an argument (empty intersections do not appear in the calculation, see [Wang 2018a]).

H-Arg a general argument type

The H-Arg might include, in theory, all argument types carrying the notions of conjunction (logical "AND") and/or disjunction (logical "OR"). Using the informal definition of the argument types given in [Cyra 2011, Ayoub 2013] and the formal one in [Wang 2019], we placed those types according to ours in figure 2.5. We can notice that each work presents a version of pure conjunctive type and a pure disjunctive one. The rest of the types can be consider as a special case of the hybrid (H-Arg) one. Some are close to the C-Arg, while others are closer to the D-Arg. The complementary and alternative combination type (closer to the H-Arg definition), proposed by Cyra and Gorski [Cyra 2011], is not closer to C-Arg than to D-Arg. However, authors decided to consider it as a alternative type or complementary according to each situation instead of proposing formulas that gather the two. We also believe that the redundant argument defined in Wang et al works has no interest since the concept of disjunction could be carried by the complementary argument type when the conjunction of premises does not bring more confidence to the conclusion $m([\wedge_{i=1}^{n} p_i] \equiv C) = \max(m(\wedge_{i=1}^{n}[p_i \equiv C]))$. In this situation, we can set $m([\wedge_{i=1}^{n} p_i] \equiv C)$ to 0, because the conjunction of elementary rules $m(p_i \equiv C)$ can express the redundancy concept (each premise can lead to the validation of the conclusion alone).

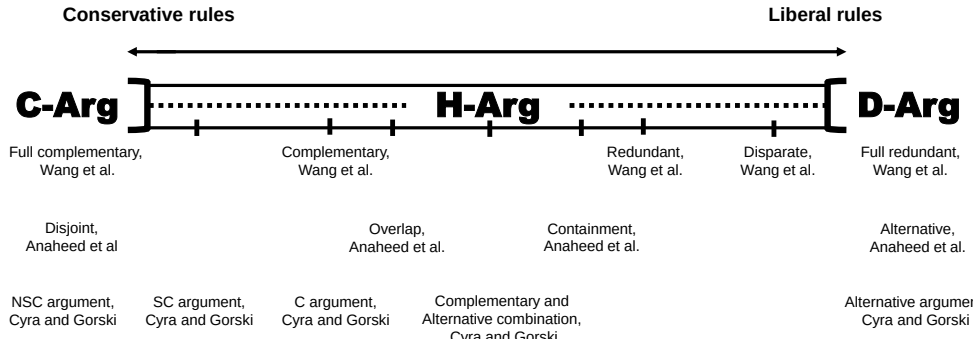


Figure 2.5: Argument types

Table 2.18: Values of the 3-tuple (Bel, $Disb$, $Uncer$) resp. for C-Arg, D-Arg and H-Arg

Values / Premise/Conclusion	Belief	Disbelief	Uncertainty
p_1	1	0	0
p_2	0	1	0
C for a C-Arg	0	1	0
C for a D-Arg	1	0	0
C for a H-Arg	0	0	1

2.5 Models sensitivity analysis

In order to illustrate the claims mentioned above about how belief, disbelief and uncertainty propagate in GSN patterns, three examples of conclusions C, each supported by two premises p_1 and p_2 each are proposed. Each one represents an argument type (C-Arg, D-Arg and H-Arg). We suppose that the mass values on rules according to each argument type (direct, reverse, conjunctive and disjunctive) are maximal (= 1). We also, assign opposite values to both premises $p_1 : Bel_p^1(p_1) = 1$ and $p_2 : Disb_p^2(p_2) = 1$ (keep in mind: $Bel^i(p_i) + Disb^i(p_i) + Uncer^i(p_i) = 1$).

Table 2.18 groups the propagation results of the 3-tuples (Bel, $Disb$, $Uncer$) for the three argument types (C-Arg, D-Arg and H-Arg) when the premises $p_1 = (1, 0, 0)$ and $p_2 = (0, 1, 0)$. We can, indeed, confirm the statements made regarding the formulas (2.2), (2.3) and (2.4). We can notice that C-Arg and D-Arg propagate (resp.) the premise with the least (minimal belief and maximal disbelief) and the greatest (maximal belief and minimal disbelief) assessment. On the other hand, we notice that the H-Arg delivers a maximal degree of uncertainty to show that it cannot take a side (resp. p_1 fully in favor or p_2 fully against of C). Having opposite information (assessments) is the same as having none, since no useful assessment about the conclusion can be inferred from it. On the other hand, notice that both C-Arg and D-Arg have conflict free values (we always have: $m_C(\perp) = 0$).

Next, we set the masses on rules and premises to new values in table 2.19. Then,

Table 2.19: Rules and premises basic values for the sensitivity analysis

Values / Premise/Rules	Belief	Disbelief	Uncertainty
$(p_1 \wedge p_2) \Rightarrow C$	1	-	0
$(\neg p_1 \wedge \neg p_2) \Rightarrow \neg C$	1	-	0
$p_1 \Rightarrow C,\ p_2 \Rightarrow C$	0.75	-	0.25
$\neg p_1 \Rightarrow \neg C,\ \neg p_2 \Rightarrow \neg C$	0.75	-	0.25
p_1	0.75	0.25	0
p_2	0.25	0.75	0

we vary one parameter at a time from its maximal value ($Bel = 1, Disb = 0$) to its minimal one ($Bel = 0, Disb = 1$), while we fix all weight of the rest to their base values. For each type, we present the sensitivity of the conclusion C (belief, disbelief and uncertainty). The results are presented in box-plots of figures 2.6 to 2.14 below. In theses figures: "Bel_dr", "Bel_dr1" and "Bel_rr2" resp. stand for belief of the conjunctive direct rule, elementary rule from the premise 1 and elementary reverse rule from the premise 2. We also present the sensitivity of $m_C(\perp)$ the conflict degree in figure 2.15.

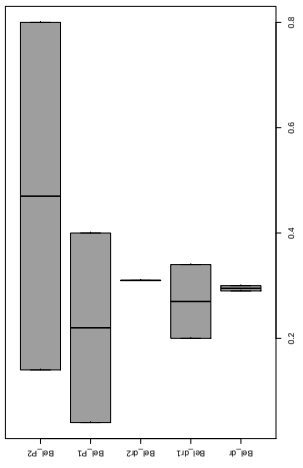


Figure 2.6: Belief sensitivity for a C-Arg

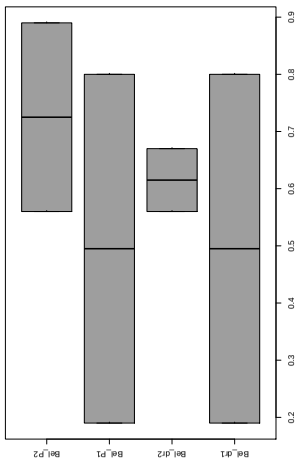


Figure 2.7: Belief sensitivity for a D-Arg

Figure 2.8: Belief sensitivity for a H-Arg

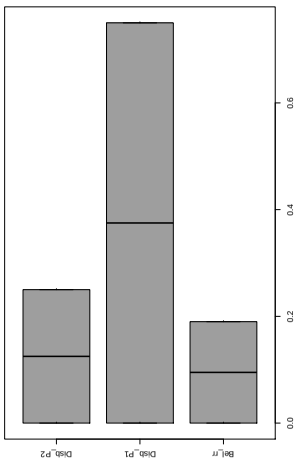


Figure 2.10: Disbelief sensitivity for a D-Arg

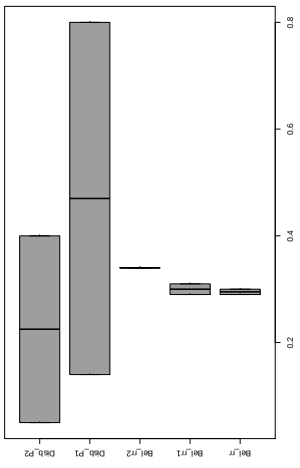


Figure 2.11: Disbelief sensitivity for a H-Arg

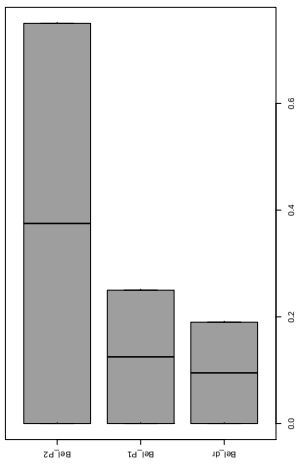


Figure 2.9: Disbelief sensitivity for a C-Arg

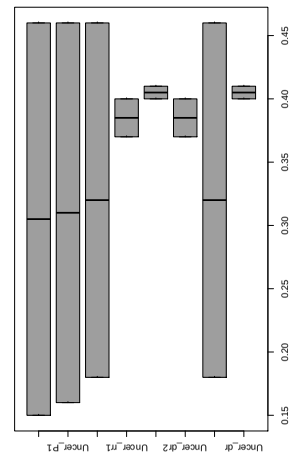


Figure 2.14: Uncertainty sensitivity for a H-Arg

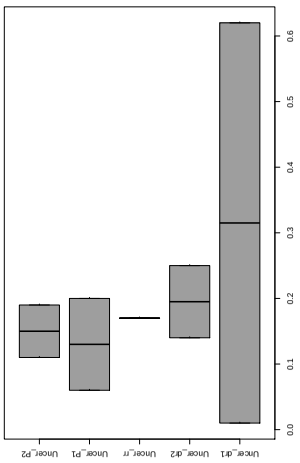


Figure 2.13: Uncertainty sensitivity for a D-Arg

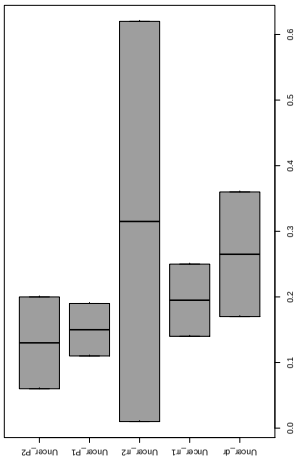


Figure 2.12: Uncertainty sensitivity for a C-Arg

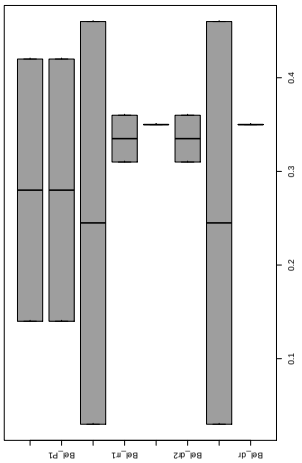


Figure 2.15: Conflict degree sensitivity

First of all, we can notice that all of belief, disbelief, uncertainty and conflict degree belong to the interval $[0, 1]$. We also find an opposite range of variability in premises between C-Arg and D-Arg. This is a first validation of the calculation of the conflict that we proposed. For instance, when the premise belief in (resp. disbelief) of p_1 has a tight (resp. large) range of variability in one argument type, the belief (resp. disbelief) on p_2 has a large (resp. tight) one in the other, and vice-versa. This result is due to the product used to calculate the belief (resp. disbelief) in C-Arg (resp. D-Arg). Hence, a tight range of values obtained. Conversely, probabilistic sum is used to calculate disbelief (resp. belief) in D-Arg (resp. C-Arg). Hence, a large range of values obtained.

We can notice that conjunctive rules (direct "Bel_dr" and reverse "Bel_rr") have no impact on conflict degree. On the other hand, disjunctive rules (direct "Bel_dr"1, "Bel_dr2" and reverse "Bel_rr1", "Bel_rr2") and the premises "Bel_p1, "Bel_p2" influence the conflict mass value. This observation is expected because only masses on rules and the disjunctive rules are involved in the calculation of $m(\perp)$, see equation 2.5.

We notice that p_1 the premise the with the highest assessment has the most impact, in the case of a D-Arg, on the conclusion C. While with C-Arg, the premise with the lowest assessment p_2 has the most influence on the conclusion C (figure 2.15). On the other hand, for the H-Arg both premises (with a slight difference in favor of p_2) influence the assessment on the conclusion. So, to improve the results, one needs to reconsider the evidence provided in p_1 for the D-Arg, p_2 for the D-Arg and both for H-Arg.

2.6 Conclusion

In this chapter we defined two parameters to quantify confidence/uncertainty in GSN patterns. The first quantifies the confidence in premises, while the second quantifies the confidence in the rule leading to the conclusion. Then, we identify four argument types that one can encounter in argument structures. To each type of argument, we associate a formula for the propagation of belief and disbelief. We notice that the so-called H-Arg represents a general frame from which we can deduce the formulas of the other types. The C-Arg and D-Arg represent limit cases of the latter (see figure 2.5).

Figure 2.16 represents the information (masses) needed to compute the confidence (belief and disbelief) in the conclusion. To get these values two pieces of information are needed. The belief values of conjunctive and elementary (disjunctive) rules and the belief and disbelief values on premises. In the next chapter we propose an approach to expert opinions elicitation, which aims to collect assessments about premises and rules in qualitative form and transform them to belief and disbelief values to feed the propagation models presented above (see figure 2.16).

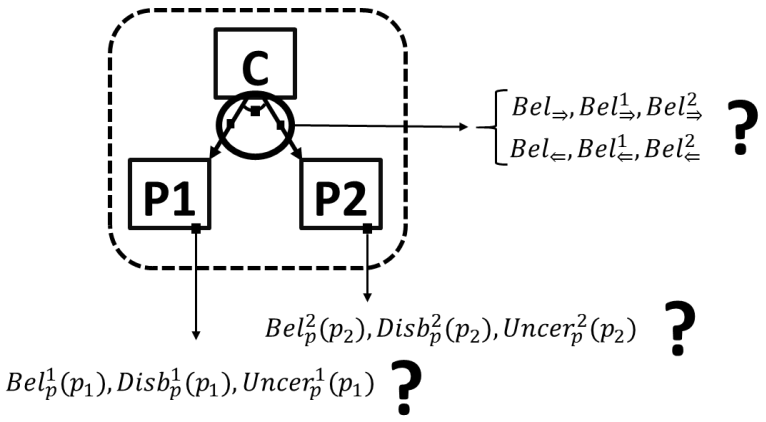


Figure 2.16: Masses on rules and premises

Chapter 3

Expert opinion elicitation

Contents

3.1 Introduction **55**
3.2 Elicitation model **56**
3.3 Belief and disbelief elicitation of premises **57**
3.4 Belief elicitation for rules **62**
3.5 Argument type identification **64**
3.6 Pignistic transform Vs Shenoy transform **65**
3.7 Conclusion **67**

3.1 Introduction

Confidence propagation models based on argument structures, particularly those based on the GSN representation, are designed on the basis of two parameters. The first represents the confidence in premises, where evidence provided may be questionable. They may be either accepted or rejected based on reasonable documented justifications (solutions in GSN formalism, see figures 1.2 and 1.3) made available to peer (expert) reviewers. For instance, an expert may doubt the validity of a test because the experimental protocol was not followed or because the sample studied is not representative. The second one represents confidence in rules (that can be assimilated to strategy component in GSN formalism). It measures the impact of the premises on the conclusion according to a defined context.

In order to use this confidence propagation model, it is necessary to provide it with the inputs to run it, i.e. the degrees of belief in premises and rules. These values will be collected from a questionnaire presented to an expert who will evaluate the argument. However, it is quite challenging, even for an expert, to give these values in term of belief and disbelief. Instead, it is more convenient to use a natural language assessment (e.g. high, medium or low confidence) and then transform it into an appropriate format (e.g., belief and disbelief degrees).

In this chapter, we present an improved version of the elicitation approach proposed by Cyra and Gorski [Cyra 2011] and reused by Wang et al. [Wang 2018a]. It consists in collecting expert opinions on premises in a qualitative form and transform them to quantitative belief (Bel_p^i) and disbelief ($Disb_p^i$) values. We also present a procedure for belief elicitation values on rules ($Bel_{\Leftarrow}, Bel_{\Rightarrow}, Bel_{\Leftarrow}^i, Bel_{\Rightarrow}^i$). Next, we

discuss the issue of argument type identification, and then compare two formulas for transforming belief functions into probability measures that can be used in the elicitation model.

3.2 Elicitation model

To evaluate a proposition (premises or rules), we choose to use two pieces of information provided by an expert:

- *Decision*: a qualitative scale which describes a trend from acceptance (maximal belief) to rejection (maximal disbelief) of a proposition. Each item of this scale (in qualitative format D_i, see figure 3.2) corresponds to a numerical value, which can be understood as the acceptance probability of a proposition x. For instance, $Dec(x) = 1$ indicates full certainty on the truth of x. On the other hand, $Dec(x) = 0$ indicates full certainty on its falsity. And, when for some reason the expert cannot take sides, $Dec(x) = 1/2$. Formally, decision is defined as the *Pignistic* transform [Smets 2005] that turns a mass function m on a set Ω_x (the frame of discernment) into a probability, changing the focal sets into uniform distributions. When $\Omega_x = \{x, \neg x\}$ has two possible states, $Dec(x)$ is the midpoint of uncertainty interval between belief and plausibility ($Pl(x) = 1 - Disb(x)$) of x.

$$Dec(x) = \frac{1 + Bel(x) - Disb(x)}{2} \tag{3.1}$$

Note that when $Bel(x) = Disb(x)$ (especially for 0), we get $Dec(x) = 1/2$.

- *Confidence*: a qualitative scale which describes the amount of information an expert possesses that can justify his/her decision. When $Conf(x) = 1$, it means that the expert has full information supporting his/her choice of $Dec(x)$. While, when $Conf(x) = 0$, it means that (s)he has no information to accept or deny (x). Formally, confidence is defined as summation of belief (evidence in favor of the proposition) and disbelief (evidence against the proposition) degrees. Confidence gives the range of the uncertainty interval : $Uncer(x) = 1 - Conf(x)$, where:

$$Conf(x) = Bel(x) + Disb(x) \tag{3.2}$$

In figure 3.1, we can notice that with both *decision* and *confidence* measures, the uncertainty interval can be defined.

Choosing scales for $(Conf, Dec)$ and translating such pairs into numerical degrees is not trivial, we thus make the equidistance assumption for simplicity (there is no apparent reason to use a particular scale) and to be comparable to previous works [Cyra 2011, Wang 2018b].

To summarize, in order to assess a proposition, an expert is asked to give his/her decision and the confidence associated to it. Both pieces of information are collected using an evaluation matrix. In figure 3.2, each dot of this matrix correspond to

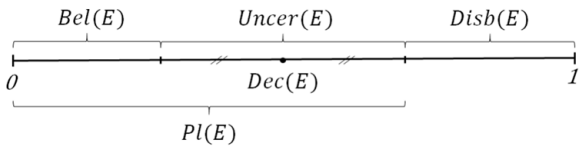


Figure 3.1: The positioning of the uncertainty interval on the pair $(Dec, Conf)$

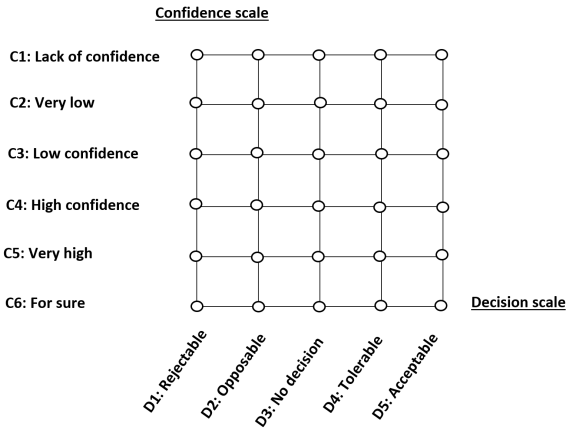


Figure 3.2: Evaluation matrix

linguistic a pair (Dec, $Conf$). The numerical values associated to each qualitative item of these two scales are used to calculate (using equations 3.1 and 3.2) a 3-tuple (Bel, $Disb$, $Uncer$) necessary to run the uncertainty propagation model defined in the previous chapter. This operation (including the reverse transformation from (Bel, $Disb$, $Uncer$) to (Dec, $Conf$)) is called *elicitation*.

In Figure 3.3, 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 3.3.a (resp. 3.3.b). It corresponds to a maximal disbelief (resp. belief). In contrast, the lower matrices represent resp. the cases of total conflict (Figure 3.3.c) and ignorance (Figure 3.3.d). In both cases, the expert cannot make a clear decision either because (s)he has as a lot of information both to support and reject the claim ($Conf = 1$), or because (s)he has no information ($Conf = 0$). In contrast to other works [Cyra 2011, Wang 2019], we allow the assessor to use a midpoint value ($Dec(A) = 1/2$) to show full hesitancy.

3.3 Belief and disbelief elicitation of premises

As presented above, each pair (Dec, $Conf$) provided by an expert is converted into a 3-tuple (Bel, $Disb$, $Uncer$) using elicitation formulas (equations 3.1 and 3.2) in order to be used in the propagation model. This means that the expert is free to choose any value from the evaluation matrix presented in the previous

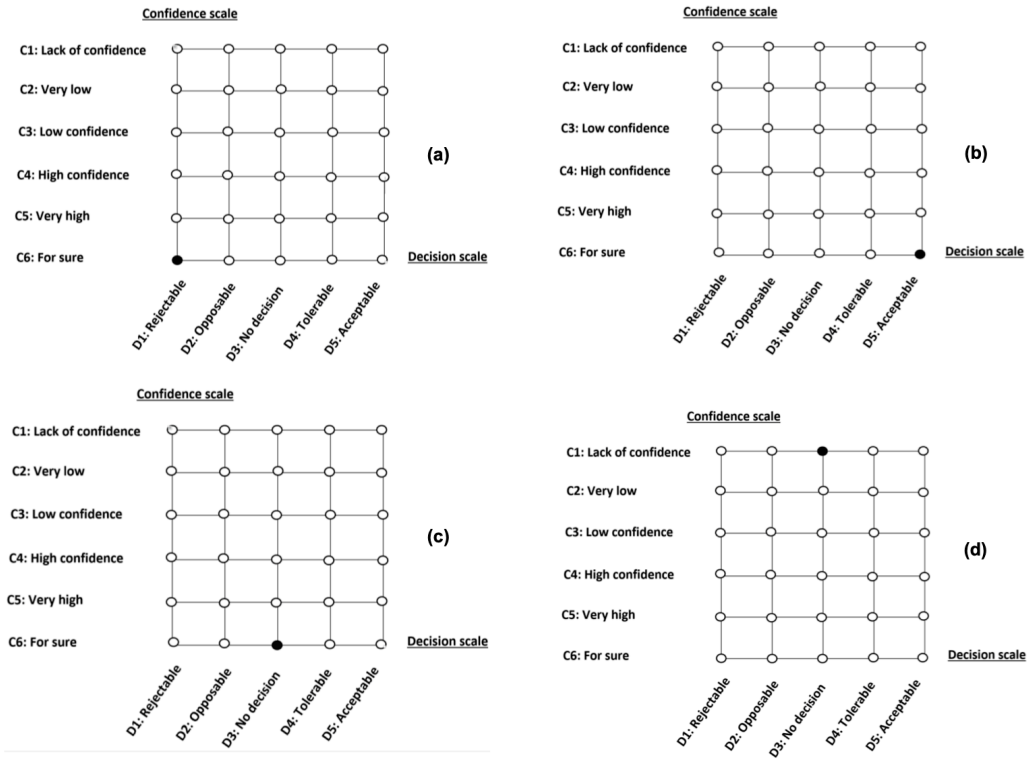


Figure 3.3: Extreme assessments (black dot)

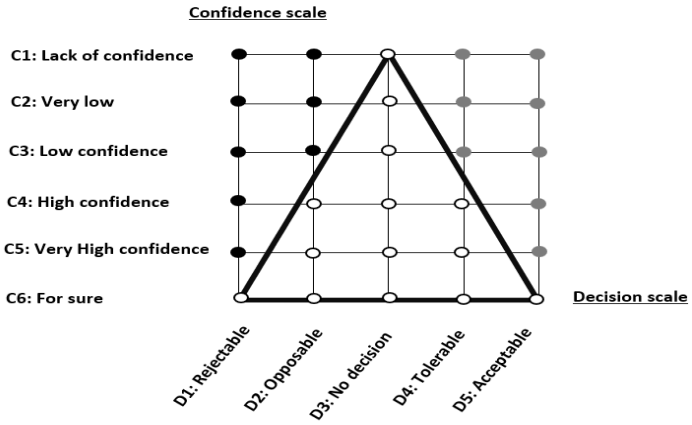


Figure 3.4: Evaluation matrix including Josang constraint

subsection. However, a constraint known as "*Josang constraint*" [Jøsang 2016] needs to be respected to get usable values. Indeed, we notice that choosing dots outside the triangle in the evaluation matrix (figure 3.4) leads to negative belief (black dots) or disbelief (grey dots) values, which make no sense. This constraint relates $Conf(p)$ and $Dec(p)$: (3.2) and (3.1) imply $1 - Conf(p) \leq \min(2Dec(p), 2(1 - Dec(p))$. To fix this problem, we can express the range of $Dec(p)$ for a given confidence level as:

$$\frac{1 - Conf(p)}{2} \leq Dec(p) \leq \frac{1 + Conf(p)}{2} \tag{3.3}$$

Proof:
There are:

$$\begin{cases} Dec(x) &= \frac{1+Bel(x)-Disb(x)}{2} \\ Uncer(x) &= 1-Bel(x)-Disb(x) \end{cases}$$

Hence, we get:

$$\begin{cases} 2\cdot Dec(x) &= 1+Bel(x)-Disb(x) \leq 2 \\ Uncer(x) &= 1-Bel(x)-Disb(x) \geq 0 \end{cases}$$

By summing and subtracting respectively these two equations,
we obtain:

$$\begin{cases} 2\cdot Dec(x)-Uncer(x) &= 2\cdot Bel(x) \geq 0, \quad (0 \leq 2\cdot Bel(x) \leq 2) \\ 2\cdot Dec(x)+Uncer(x) &= 2\cdot[1-Disb(x)] \leq 2, \ (0 \leq 2\cdot[1-Disb(x)] \leq 2) \end{cases}$$

So,

$$\begin{cases} Uncer(x) \leq 2\cdot Dec(x) \\ Uncer(x) \leq 2\cdot(1-Dec(x)) \end{cases}$$

We therefore conclude that:

$$Uncer(x) \leq \min(2\cdot Dec(p), 2\cdot(1-Dec(p))$$

When an expert makes a strong decision (acceptance with $Dec(p) = 1$, or rejection with $Dec(p) = 0$) the confidence must be maximal ($Conf(p) = 1$), otherwise his/her assessment will have no grounds. The closer you get to the midpoint value ($Dec(p) = 1/2$, no decision), the larger the confidence interval can be. ($Dec = 1/2$, $Conf = 0$) means that the expert cannot take a side because he has no information (*total ignorance*), while ($Dec = 1/2$, $Conf = 1$) means that he/she cannot take a side because he/she has as much evidence in favor of the premise as against (*total conflict*). So, when the pair $(Dec(p), Conf(p))$ is situated outside the triangle, and $Dec(p) < \frac{1-Conf(p)}{2}$ (rejection: black dots on figure 3.2), we set $Dec(p) = \frac{1-Conf(p)}{2}$. On the other hand, when $Dec(p) > \frac{1+Conf(p)}{2}$ (acceptance: grey dots on figure 3.2), we set $Dec(p) = \frac{1+Conf(p)}{2}$.

Example 1. Suppose a conclusion (C) supported with two premises (p_1) and (p_2). Each of these premises is associated with one or more pieces of evidence (Sn_i). Figure 3.5 summarizes the elicitation procedure of belief and disbelief of these premises. Firstly, based on the evidence provided for each premise the expert needs to give his/her assessment of each premise by choosing a pair (Dec,Conf), using the evaluation matrices. Assuming the expert gives the assessments below:

- p_1: Opposable with high confidence ($Dec(p_1) = 0.25, Conf(p_1) = 0.6$).

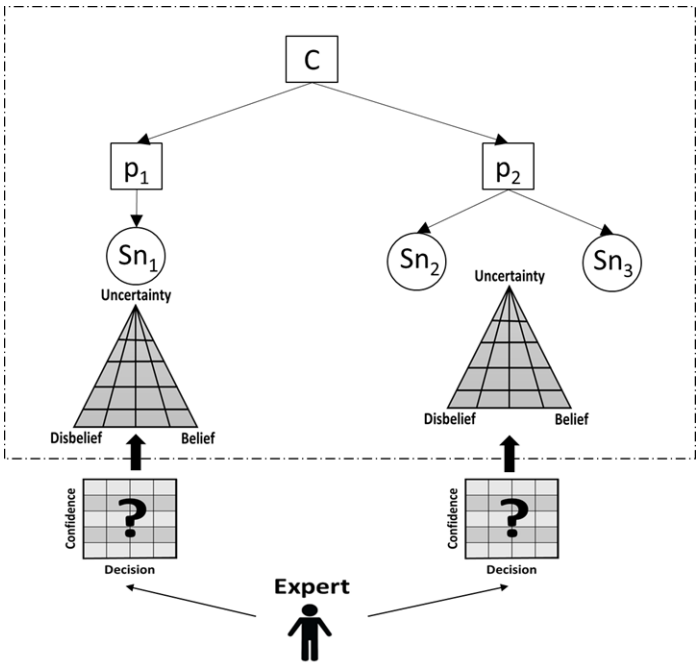


Figure 3.5: Belief and disbelief elicitation of premises

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

The second step is to calculate $Bel(p_i)$ and $Disb(p_i)$. To do this, we use the numerical values $Dec(p_i)$ and $Conf(p_i)$ associated to each assessment and the following formulas: $Bel(p) = \frac{Conf(p)-1}{2} + Dec(p)$, $Disb(p) = \frac{Conf(p)+1}{2} - Dec(p)$ (deduced from equations 3.1 and 3.2).

We can notice that the assessment for p_1 is inside the triangle in the matrix (figure 3.2). We can therefore claim that 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 $Uncer(p_1) = 1 - Bel(p_1) - Disb(p_1) = 0.4$.

On the other hand, the assessment for p_2 is situated outside the triangle. In this case, we can be sure that 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 $Uncer(p_2) = 1 - Bel(p_2) - Disb(p_2) = 0.2$.

In table 3.1, we grouped all possible pairs (Dec, $Conf$) on premises with their appropriate (Bel, $Disb$) counterparts. For this first work we stick to the chosen values for $(Dec, Conf)$ in the previous subsection. The degree of uncertainty can easily be deduced ($Uncer(p) = 1 - Bel(p) - Disb(p)$). We can notice an anti-symmetry between belief and disbelief degree regarding the central column (D3: No decision). We also notice that when no information is available (C1: Lack of

Table 3.1: Values from (Dec, $Conf$) to (Bel,$Disb$) couples on premises (see figure 3.2 for symbol meaning)

$Conf$ \ Dec	D1 (0)	D2 (0.25)	D3 (0.5)	D4 (0.75)	D5 (1)
C1 (0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)
C2 (0.2)	(0,0.20)	(0,0.20)	(0.10,0.10)	(0.20,0)	(0.20,0)
C3 (0.4)	(0,0.40)	(0,0.40)	(0.20,0.20)	(0.40,0)	(0.40,0)
C4 (0.6)	(0,0.60)	(0.05,0.55)	(0.30,0.30)	(0.55,0.05)	(0.60,0)
C5 (0.8)	(0,0.80)	(0.15,0.65)	(0.40,0.40)	(0.65,0.15)	(0.80,0)
C6 (1)	(0,1)	(0.25,0.75)	(0.50,0.50)	(0.75,0.25)	(1,0)

confidence), no matter what choice is made the degree of uncertainty is maximal ($Uncer(p) = 1$). On the other hand, in the case of a fully informed expert (C6: For sure) the decision value varies from rejection to acceptance, which follows the restrictions imposed by "Josang constraint".

3.4 Belief elicitation for rules

In the previous subsection, we presented how to estimate $(Bel, Disb, Uncer)$ of a premise. In this subsection we study how to estimate $(Bel, Uncer)$ for the rules (keep in mind that disbelief on rules is not defined, see chapter 2) in order to propagate the confidence in the GSN.

Before presenting the details of the belief estimation procedure of the rules, we draw your attention to an observation made on the propagation model presented in chapter 2. The whole procedure is based on this observation. 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 ($Uncer^i_p(p_i) = 1$), $Bel_C(C)$ and $Disb_C(C)$ can take (resp.) the belief values of direct and reverse rules. For example, 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, i.e. $Bel^1_p(p_1) = Bel^2_p(p_2) = 1$ then: $Bel_C(C) = Bel_{\Rightarrow}([p_1 \wedge p_2] \Rightarrow C)$ using equation (2.2). While assuming total reject with maximal confidence of (p_1), i.e $Disb^1_p(p_1) = 1$ and total ignorance about (p_2), i.e. $Uncer^2_p(p_2) = 1$ then: $Disb_C(C) = Bel_{\Leftarrow}(\neg p_1 \Rightarrow \neg C)$ using equation (2.2).

We propose a procedure for collecting belief of rules based on the elicitation model (as for the premises) and the observation of the paragraph above where we assume that the GSN pattern to be assessed is a C-Arg (resp. D-Arg) to estimate the values of $Bel_{\Rightarrow}([\wedge^n_{i=1} p_i] \Rightarrow C)$ and $Bel^i_{\Leftarrow}(\neg p_i \Rightarrow \neg C)$ (resp. $Bel^i_{\Rightarrow}(p_i \Rightarrow C)$ and $Bel_{\Leftarrow}([\wedge^n_{i=1} \neg p_i] \Rightarrow \neg C)$). Thus, from the expert assessment of the conclusion for predefined premises assessments $(Bel, Disb, Uncer)$, we conclude the belief values of the rules. These values can be used afterwards to calculate the confidence of a conclusion for all values of the premises.

Moreover, as presented in subsection 2.2, the disbelief of the rule is not defined

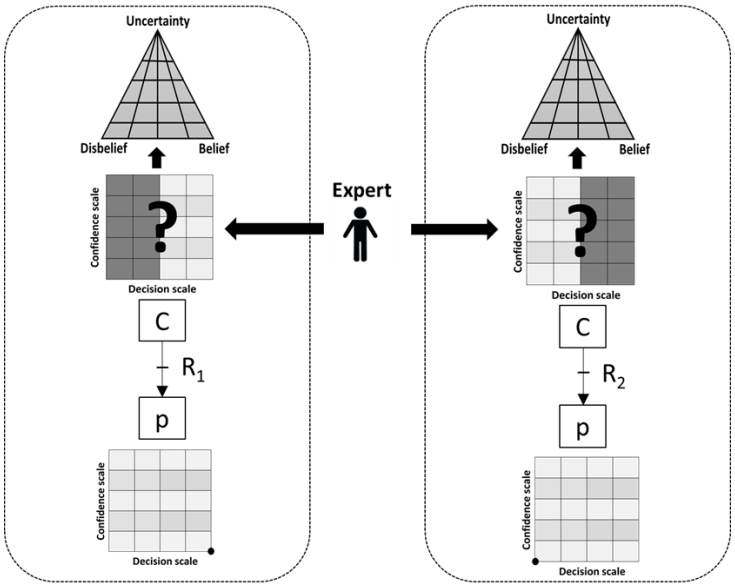


Figure 3.6: Belief elicitation of rules

which is the same as assuming that it has a value equal to zero. 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 negative decision (from "*rejectable*" to "*no decision*") for the reverse rules. Formulas in (3.2) and (3.1) are used to derive the degrees of belief on rules.

Example 2. Suppose the case of a conclusion (C) supported by one premise (p). Figure 3.6 describes the procedure of rules elicitation. To get the belief on its direct rule $R_1 : p \Rightarrow C$ and reverse one $R_2 : \neg p \Rightarrow \neg C$, we ask an expert to give his/her assessment about the conclusion respectively when $Bel_p(p) = 1$ (which is equivalent to $(Dec(p) = 1,\ Conf(p) = 1)$ for R_1, and then when $Disb_p(p) = 1$ (which is equivalent to $(Dec(p) = 0,\ Conf(p) = 1)$ for R_2. Suppose the expert gives the following assessments:

- Direct rule $(R_1 : p \Rightarrow C)$: Assuming $(Dec(p) = 1,\ Conf(p) = 1)$, the expert assigns "Tolerable, with high confidence" to the conclusion (C): $Dec(C) = 0.75,\ Conf(C) = 0.6$

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

We can notice in this example that both cases respect the Josang constraint (3.3). Hence, there is no need to adjust the decision value. Using equations (3.2) and (3.1) for the direct rule R_1: $Bel_{\Rightarrow}(R_1) = Bel_C(C) = \frac{(0.6)-1}{2} + (0.75) = 0.55$

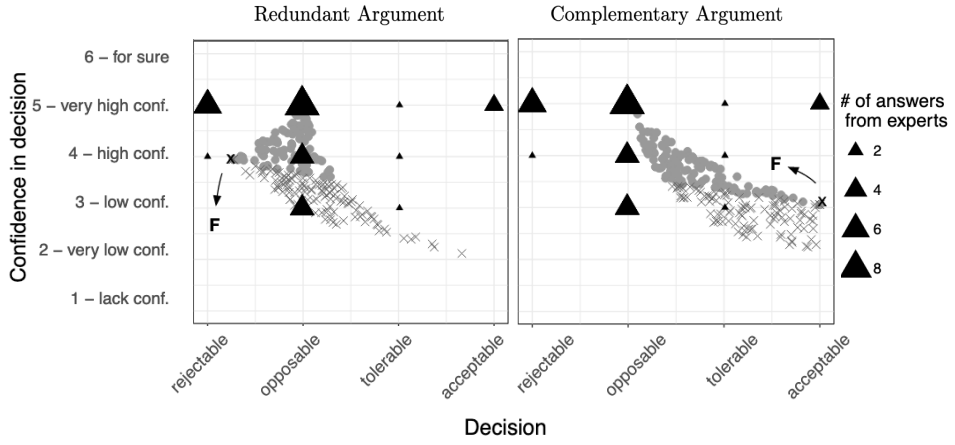


Figure 3.7: Responses from experts vs. Calculated results based on Complementary and Redundant argument types [Wang 2018b]

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 $Bel_{\Leftarrow}(\neg R_2) = 0$.

We note that the results obtained for these two rules are consistent with our expectations. Indeed, starting from high confidence values, we find that the belief value of these rules is indeed higher than the uncertainty value.

3.5 Argument type identification

In our approach H-Arg encompasses C-Arg and D-Arg, but it might be useful to identify if we have an extreme case (C-Arg or D-Arg). Wang et al. [Wang 2018b] use a theoretical data plot of all possible values of the masses on the rules (appropriateness ω_i), where they assume that the assessed argument is a complementary one and then a redundant one (the main types defined in their work, the others represent special cases of these two), and cross them with the data collected from the experts. The chosen argument type is the one with the most intersection between expert opinion and theoretical data, see figure 3.7.

In our case, this step is not required since we propose a general formula that propagates the confidence in the three types we defined. Remember that the conjunctive (C-Arg) and disjunctive (D-Arg) argument types represent the limit cases of a general case that we named hybrid argument type (H-Arg). Thus, it is enough

to observe the values of the elicited rules to know if the studied GSN pattern represents an extreme case (D-Arg or C-Arg) or not. Indeed, let us recall that if all direct elementary rules ($p_i \Rightarrow C$) and the reverse conjunctive one ($[\wedge_{i=1}^{n} \neg p_i] \Rightarrow \neg C$) take the value zero for GSN pattern, we can deduce that this studied GSN pattern corresponds to a C-Arg. Conversely, if the conjunctive rule $[\wedge_{i=1}^{n} p_i] \Rightarrow C$) and reverse elementary rules ($\neg p_i \Rightarrow \neg C$) take the value zero, we can say that the pattern corresponds to a D-arg. To maintain better consistency with the definition of D-Arg, the conjunctive rule is set two zero when it does not bring additional confidence to the conclusion ($Bel_{\Rightarrow}([\wedge_{i=1}^{n} p_i] \Rightarrow C) = \max(Bel_{\Rightarrow}^{i}(p_i \Rightarrow C))$. Apart from these two cases, we consider that the pattern corresponds to an H-Arg.

To summarize, here are the rules for argument type identification:

- If $Bel_{\Rightarrow}^{i}(p_i \Rightarrow C) = 0$, GSN $\mapsto$ C-Arg.
- If $\max(Bel_{\Rightarrow}^{i}(p_i \Rightarrow C)) = Bel_{\Rightarrow}([\wedge_{i=1}^{n} p_i] \Rightarrow C)$, $Bel_{\Rightarrow}([\wedge_{i=1}^{n} p_i] \Rightarrow C) = 0$ and GSN $\mapsto$ D-Arg.
- If $Bel_{\Rightarrow}([\wedge_{i=1}^{n} p_i] \Rightarrow C) > \max(Bel_{\Rightarrow}^{i}(p_i \Rightarrow C)) > 0$, GSN $\mapsto$ H-Arg.

Since the argument type is based on the values taken by the beliefs on the direct rules and that the inverse rules are deduced from the direct ones, we expect that when $Bel_{\Rightarrow}^{i}(p_i \Rightarrow C) = 0$, then $Bel_{\Leftarrow}([\wedge_{i=1}^{n} \neg p_i] \Rightarrow \neg C) = 0$ for the C-Arg; and when $Bel_{\Rightarrow}([\wedge_{i=1}^{n} p_i] \Rightarrow C) = 0$, then $Bel_{\Leftarrow}^{i}(\neg p_i \Rightarrow \neg C) = 0$ for the D-Arg.

3.6 Pignistic transform Vs Shenoy transform

In this section, we study an alternative to our decision formula 3.1, which uses the *Pignistic* transform, to make a comparison.

In the literature, there are formulas other than the *Pignistic* transform that transforms degrees of belief, degrees of possibility, etc. into probabilities Among these, an interesting alternative presented in chapter 1 known as *Shenoy* [Cobb 2006] transform can be used to define the decision parameter.

Using Shenoy transform on a two-state frame of discernment $\Omega_x = \{x, \neg x\}$, we get the following decision formula:

$$Dec(x) = \frac{Pl(x)}{Pl(x) + Pl(\neg x)} = \frac{1 - Disb(x)}{2 - (Bel(x) + Disb(x))} \tag{3.4}$$

The confidence formula remains the same $Conf(x) = Bel(x) + Disb(x)$. From these two equations ($Conf$ and Dec), we deduce belief and disbelief expression:

$$\begin{cases} Disb(x) &= 1 - Dec(x) \cdot [2 - Conf(x)] \\ Bel(x) &= Conf(x) - Disb(x) \end{cases} \tag{3.5}$$

To avoid getting negative belief and disbelief degrees for some decision and confidence values a constraint that frames each decision between two confidence values

needs to be respected.

$$\frac{1}{2 - Conf(x)} \leq Dec(x) \leq \frac{1 - Conf(x)}{2 - Conf(x)} \tag{3.6}$$

Proof:
There are:

$$\begin{cases} Disb(x) &= 1 - Dec(x) \cdot [2 - Conf(x)] \geq 0 \\ Bel(x) &= Conf(x) - Disb(x) \geq 0 \end{cases}$$

By replacing $Disb$ by its expression in the second equation, we obtain:

$$\begin{cases} Disb(x) &= 1 - Dec(x) \cdot [2 - Conf(x)] \geq 0 \\ Bel(x) &= Conf(x) - 1 + Dec(x) \cdot [2 - Conf(x)] \geq 0 \end{cases}$$

Hence, we get $(\forall x \in [0,1], Conf(x) \leq 1)$:

$$\begin{cases} Dec(x) &\leq \frac{1}{2-Conf(x)} \\ Dec(x) &\geq \frac{1-Conf}{2-Conf(x)} \end{cases}$$

Figure 3.8 presents the aeras from which transformation formulas (triangle for Pignistic 3.3 and curvy inner triangle for Shenoy 3.6) give negative belief and disbelief values (before adjustment).

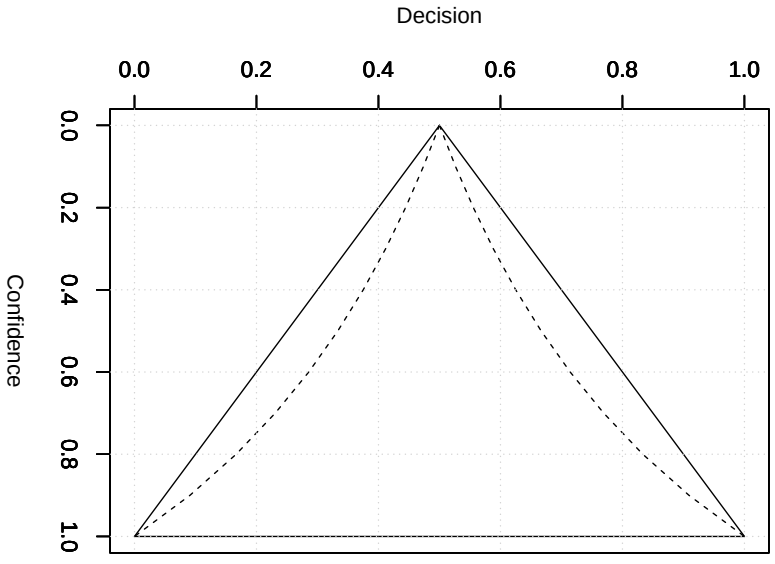


Figure 3.8: Pignistic (plain line) and Shenoy (dashed line) transforms constraint

We notice that the usable values (not requiring adjustment) represented by the

area inside the two triangles is larger in the case where we use the Pignistic transform (triangle with plain edges) than in the Shenoy case (curvy triangle with dashed edges). This observation represents a first argument in favor of using the Pignistic transform. The second argument in favor of the Pignistic transform, consists in what it brings as information. Indeed with the confidence formula, it gives, respectively, the midpoint value of the uncertainty interval and its length which is natural, but the Shenoy transform does not.

3.7 Conclusion

In this chapter, we defined two parameters used to express assessments of experts regarding a GSN pattern in natural language, and then transform them to belief and disbelief values required to compute the uncertainty propagation model proposed in chapter 2. Each assessment is collected in terms of decision about a proposal (from reject to accept) and degree of confidence the expert can associate to it (from lack of confidence, to for sure). To do so, we provided a new elicitation model that takes into consideration the situations of total ignorance, where in the previous works an arbitrary decision value is given. It also considers the case of indecision by introducing a neutral item ("no decision") in the decision scale which is very important in the evaluation process of the rules. Unlike previous work, we also use our elicitation model to collect beliefs of rules.

This elicitation method is very important in the process of assessing confidence in an argument because it provides all the inputs needed to use the propagation model. Thus the following chapter will a present complete confidence assessment procedure that integrates both propagation and elicitation models. Followed by an application on a case study.

Chapter 4

Applying confidence assessment of an argument in GSN - Case study

Contents

4.1 Introduction 69
4.2 Confidence assessment procedure 70
4.2.1 Modelling phase: Rules elicitation 70
4.2.2 Application phase: Premises elicitation and confidence propagation 73
4.3 Case study 75
4.4 Discussion 81
4.5 Conclusion 82

4.1 Introduction

Building an uncertainty/confidence assessment procedure in GSN pattern requires reliable methods of elicitation and propagation. The elicitation step should produce inputs in usable format (belief and disbelief degrees) for the propagation step. The latter should propagate uncertainties related to premises to the conclusion, while respecting the specificity of each argument structure (how premises interact to support the conclusion). Then the results should be expressed in a meaningful format for users.

In the previous chapter, we proposed an approach that

(i) collects uncertainty assessments about premises and rules, appearing in the GSN structure. This is done in a qualitative format, in terms of decisions for or against the acceptance of the involved statements (e.g., high level safety requirements coverage is achieved) and the degree of confidence justifying each of them.

(ii) expresses the qualitative data in a quantitative format then translated into belief and disbelief degrees.

We also proposed an approach (presented in chapter 2) that takes elicitation results (belief and disbelief) and propagate them to the conclusion.

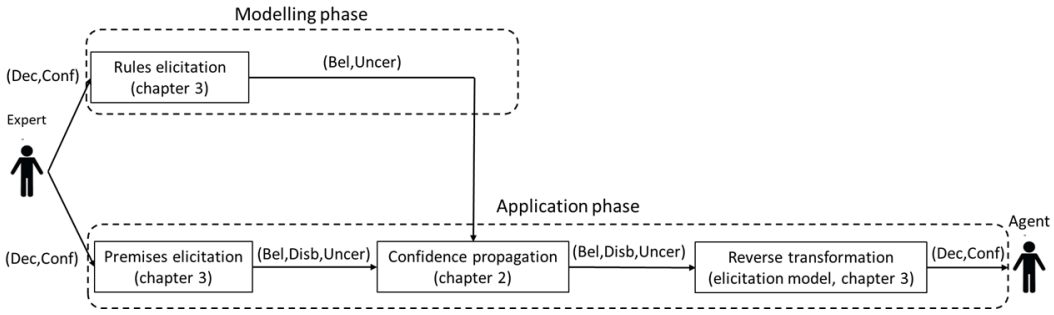


Figure 4.1: Uncertainty assessment procedure

In this chapter, we detail this procedure. We also present a case study to which we apply this approach and raise any possible issues to be addressed, see its limitations and finally validate it.

4.2 Confidence assessment procedure

Figure 4.1 summarizes the evaluation procedure structured in two phases. The first one is called *modeling phase*. It provides the beliefs on the rules. The second is called the *application phase*. It provides the beliefs on the premises and then propagates them with the beliefs on the rules, using the propagation model, up to the conclusion.

4.2.1 Modelling phase: Rules elicitation

The modeling phase collects expert opinions on rules, expressed with qualitative scores $(Dec, Conf)$, and translates them into numerical mass assignments to rules. Figure 4.2 illustrate this phase.

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 (resp.) the values of $Bel^1_{\Leftarrow}(\neg p_1 \Rightarrow \neg C)$ and $Bel^1_{\Rightarrow}(p_1 \Rightarrow C)$ the expert will be asked the following questions (in the case $n = 2$):

1. Supposing no knowledge about the premise p_2: $(Dec = 0.5, Conf = 0)$ and total reject (rejectable for sure) of p_1: $(Dec = 0, Conf = 1)$, what is your Decision/Confidence in the conclusion (see figure 4.3)?

2. Supposing no knowledge about the premise p_2: $(Dec = 0.5, Conf = 0)$ and total acceptance (acceptable for sure) $p_1 : (Dec = 1, Conf = 1)$, what is your Decision/Confidence in the conclusion (see figure 4.4)?

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

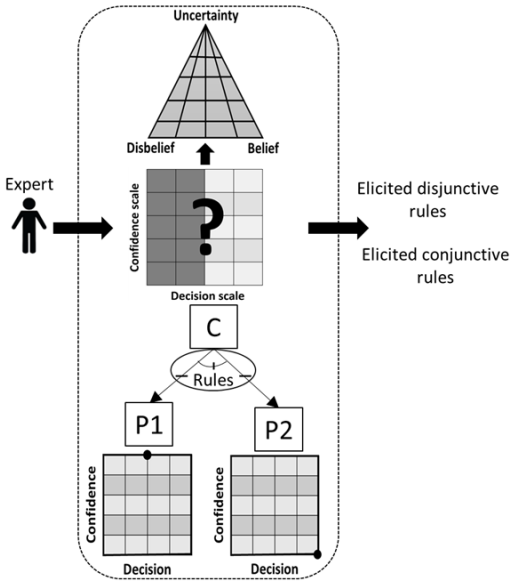


Figure 4.2: Modelling phase

3. Supposing total reject of both premises $p_1, p_2 : (Dec = 0, Conf = 1)$, what is your Decision/Confidence in the conclusion (see figure 4.5)?

4. Supposing total acceptance of both premises $p_1, p_2 : (Dec = 1, Conf = 1)$, what is your Decision/Confidence in the conclusion (see figure 4.6) ?

Figure 4.3: Disjunctive reverse rule elicitation

Figure 4.4: Disjunctive direct rule elicitation

Figure 4.5: Conjunctive reverse rule elicitation

Figure 4.6: Conjunctive direct rule elicitation

It is important to mention that the expert can only select pairs $(Dec, Conf)$ from the positive side of the evaluation matrix ("no decision" to "acceptable") while assessing direct rules. Conversely, he/she can only select negative assessment for the reverse rules ("rejectable" to "no decision"). For instance in figure 4.4, choosing a pair from this forbidden zone (shaded) will assign a mass to $Bel^1_{\Rightarrow}(\neg[p_1 \Rightarrow C])$ which is not a rule (see, previous chapters). Even if we set all the positive masses of this kind (disbelief on rules) to zero, we prefer not to allow the expert access to this area to keep consistency with the rules definition. The implication used to define rules can only infer one side of the assessment at a time.

Once the masses on rules are obtained, one can deduce the argument type of the assessed GSN pattern (C-Arg, D-Arg or H-Arg). The case of C-Arg is simple to identify, since verification of one premise (i.e., true) cannot infer the conclusion C alone. Thus, $Bel^i_{\Rightarrow}(p_i \Rightarrow C) = 0$. Conversely, the denial of one premise (i.e., false) infers the conclusion denial $\neg C$. However, it is not the same for D-Arg. If a premise p_1 supports the conclusion C, then $p_1 \wedge p_2$ also supports it even if p_2 cannot. To keep consistency with the definition of D-Arg, we set the mass of the direct conjunction and the reverse elementary ones to zero ($Bel_{\Rightarrow}([\wedge^n_{i=1} p_i] \Rightarrow C) = Bel^i_{\Leftarrow}(\neg p_i \Rightarrow \neg C) = 0$), if at least the mass on one elementary rule is equal to the one on the conjunction. In this case, we can say that the conjunction of premises does not bring additional support to the conclusion.

Then, once the masses of the rules are acquired and the propagation formula is specified, we can proceed to the next step of this assessment procedure by following the instructions below, for the considered system.

4.2.2 Application phase: Premises elicitation and confidence propagation

The application phase concerns the collection of expert data on premises and propagate them, along with data on rules, to the conclusion. Figure 4.7 illustrate this phase.

One question per premise is then formulated to the experts to collect data about premises:

- Considering the knowledge on the pieces of evidence (also called *solutions* in GSN formalism), what is your "Decision" and "Confidence" in the premise p_i (see figure 4.8)?

Grouped in a questionnaire, these $(3n+2)$ questions, the assessor will be asked to choose a pair from the evaluation matrix (some matrices being pre-filled, for rules, see Figure 4.7). Then, these values (on rules and premises) are used to calculate the belief/disbelief in the conclusion (equations **??**).

This procedure will be iterated for each node consisting of a conclusion (goal in GSN formalism) supported by its premises (sub-goals), starting from the bottom of the GSN, up to the top goal.

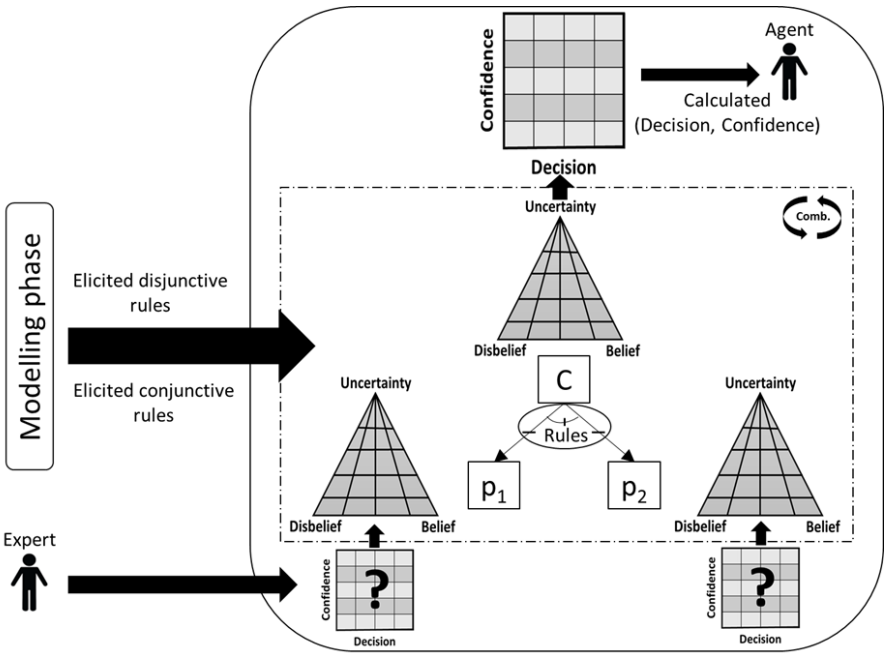


Figure 4.7: Application phase

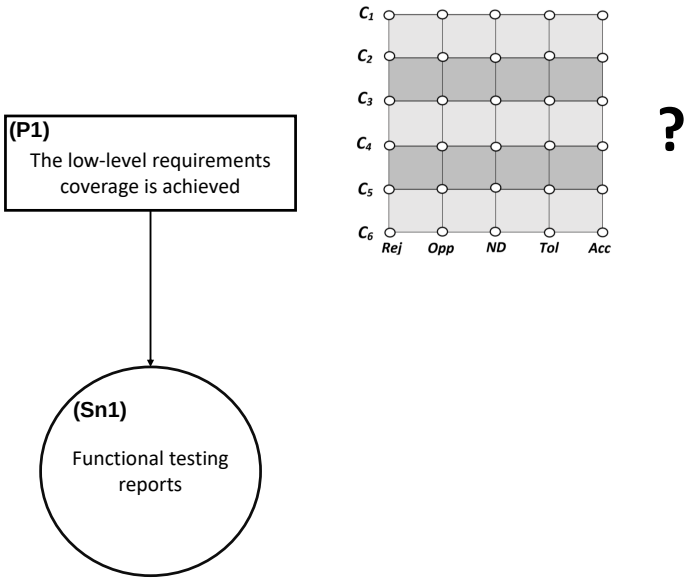


Figure 4.8: Premises elicitation

Table 4.1: Decision items and their appropriate quantitative counterpart

Dec	Qual.	*Rej* (D_1)	*Opp* (D_2)	*ND* (D_3)	*Tol* (D_4)	*Acc* (D_5)
	Qnt.	$[0, 0.125[$	$[0.125, 0.375[$	$[0.375, 0.625[$	$[0.625, 0.875[$	$[0.875, 1[$

Table 4.2: Confidence items and their appropriate quantitative counterpart (see figure 3.2 for symbols meaning)

$Conf$	Qual.	C_1	C_2	C_3	C_4	C_5	C_6
	Qnt.	$[0, 0.1[$	$[0.1, 0.3[$	$[0.3, 0.5[$	$[0.5, 0.7[$	$[0.7, 0.9[$	$[0.9, 1[$

Finally, we may transform the resulting 3-tuple (Belief, Disbelief, Uncertainty), of the conclusion, to a pair (Decision, Confidence) using formulas (3.2) and (3.1) and approximate them by choosing the qualitative values, of the closest pair (Dec,Conf) to their corresponding numerical values. To simplify, tables 4.1 and 4.2 present (resp.) the corresponding qualitative $(Dec, Conf)$ values for the conclusion outcome interval. For instance, If $Dec = 0.3$, the decision will be "Opposable". The value $Dec = 0.3$ is closer to the decision "Opposable", for which the corresponding numerical value is (0.25), than "no decision" (0.5).

4.3 Case study

In this section, we use a part of an argument proposed in [Damour 2021] to test and validate our confidence propagation approach. This study proposed an architecture of a collision avoidance system for drones, Urban Air Mobility and present a complete argument (assurance case) of it. The studied system 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 (v_{own}, v_{int}), their relative positions (θ, ψ and ρ), the time until the loss of vertical separation occurs (τ) and the previous advisory (pa), see figures 4.9 and 4.10) by a neural network (NN) of much smaller size. From five different maneuvers (CoC: clear of conflict, WL: weak left, SL: strong left, WR: weak right and SR; strong right), the system delivers the one with the least cost (i.e, probability to have a collision).

In addition to the NN-based controller, this architecture (figure 4.9) 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 authors built an assurance case for the top goal " Real-world situations where the MLM (Machine learning Model) is not robust are identified and mitigated". The strategy is to agree that the proposed system will always detect situations when the NN cannot perform its task with full guaranties and will manage it. They use a graphical adaptation of Toulmin notation proposed by the RESSAC project. This notation shares many features with the GSN. Thus, it is possible to switch from

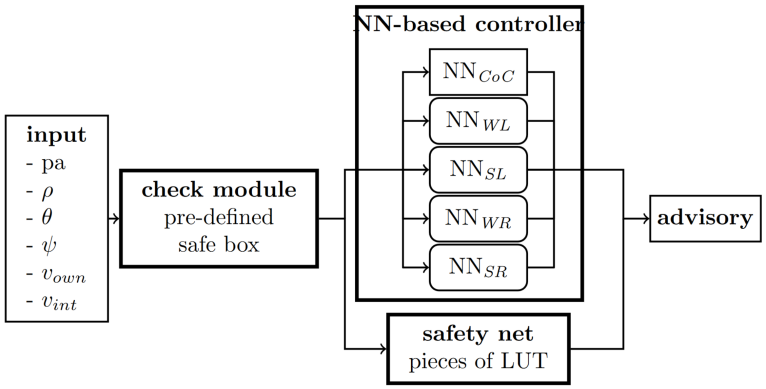


Figure 4.9: Architecture of the neural network based ACAS Xu [Damour 2021]

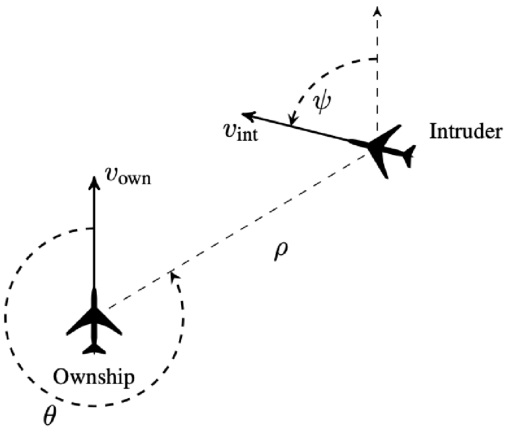


Figure 4.10: ACAS Xu geometry [Katz 2017]

one notation to the other. Figure 4.11 presents the assurance case that we encoded with GSN formalism in order to apply to it our confidence assessment method. To demonstrate this claim, 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 situations (G_4) and formally checked (G_6) in each of the areas (noted "l_i" and called p-boxes) into which the input space was correctly decomposed (G_5). This property is satisfied when the set of reachable advisories (CoC, SL, WL, WR, SR) of the NN is included in the set of reachable advisories of the LUT, i.e:

$$decisions\ NN(l_i) \subseteq decisions\ LUT(l_i)$$

We consider, in this situation, that a NN behaves similarly to the LUT on a p-box (l_i) .

Figure 4.11: Assurance Case - ML subsystem robustness [Damour 2021]

Table 4.3: Elicited belief degrees on rules

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

In order to apply our method to this case study, we worked with two authors of this work (Gabreau, C. and Pagetti, C) to validate the GSN and to ask them as experts to provide values for assessment. This case study is mainly a proof of concept, because of the small number of experts, and also because they were also the developers of the system.

Modeling phase

Following the assessment procedure above, we use a questionnaire to ask safety experts to give their opinions on the system, see Appendix section A. Table 4.3 groups the degrees of belief on the rules involved in this case. As stated in chapters above, the belief in the conjunctive reverse rule (in the form of $Bel_{\Leftarrow}([\wedge_{i=1}^n \neg p_i] \Rightarrow C)$) were set to zero (see the table), because they do not provide more confidence to their goals (G_1 and G_2). 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 involves two conjunctive argument types (arguments 1 and 2 in figure 4.12) where all premises are needed to support the conclusion and a simple one (argument 3 in figure 4.12).

Application phase

We also asked the expert, using the same questionnaire, to give their opinion about premises. The results are displayed in table 4.4.

As seen in chapter 2, 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), see table 4.4 for the appropriate 3-tuple $(Bel, Disb, Uncer)$ corresponding values, to a tolerable top goal (G_1) with low confidence ($Dec = \frac{1+0.384-0}{2} = 0.692, Conf = 0.384 + 0 = 0.384$, see table 4.5 for belief and disbelief calculation details).

Table 4.4: Basic assessment values of premises $(Bel_p(G_i), Disb_p(G_i), Uncer_p(G_i))$

Goal	G_4	G_5	G_6	G_7
Value	(1,0,0)	(0.6,0,0.4)	(0.8,0,0.2)	(0.8,0,0.2)

Table 4.5: Belief and disbelief degrees of (sub-)goals calculation

Belief, disbelief degrees
$Bel_C(G_2) = Bel_p(G_4) \cdot Bel_p(G_5) \cdot Bel_p(G_6) = 1 \times 0.6 \times 0.8 = 0.48$ $Disb_C(G_2) = 1 - [1 - Disb_p(G_4)] \cdot [1 - Disb_p(G_5)] \cdot [1 - Disb_p(G_6)] = 0$
$Bel_C(G_3) = Bel_p(G_7) = 0.8$ $Disb_C(G_3) = Disb_p(G_7) = 0$
$Bel_C(G_1) = Bel_p(G_2) \cdot Bel_p(G_3) = 0.48 \times 0.8 = 0.384$ $Disb_C(G_1) = 1 - [1 - Disb_p(G_2)] \cdot [1 - Disb_p(G_3)] = 1 - 1 = 0$

Sensitivity analysis

Graphs in figures 4.13 and 4.14 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 assessment of a premise from its minimal $(Disb_p(G_i) = 1)$ to its maximal $(Bel_p(G_i) = 1)$ 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"* $(Dec = 0, Conf = 1)$, when $Bel_p(G_i) = 0,\ i = \{4, 5, 6, 7\}$, to *"Tolerable with high confidence"* $(Dec = 0.82, Conf = 0.64)$, when $Bel_p(G_5) = 1$. 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).

Figure 4.12: Argument fragments according to their type

Figure 4.13: Decision sensitivity on the top goal G_1

Figure 4.14: Confidence sensitivity on the top goal G_1



Figure 4.15: Impact of changing the scale - example

Table 4.6: Impact of changing the decision scale on confidence assessment

Dec. scale \ Argument	C-Arg	D-Arg
$\|Dec\| = 4$	$Dec = 0.4489$, $Conf = 1$ Opposable, for sure	$Dec = 0.5511$, $Conf = 1$ Tolerable, for sure
$\|Dec\| = 5$	$Dec = 0.5625$ No decision, for sure	$Dec = 0.4375$ No decision, for sure

4.4 Discussion

We can notice from this case study that both decision and the degree of confidence associated to it strongly depend on the choice of the numerical scale used. Indeed, changing the scale could lead to different results, particularly when we approximate the numerical value of the pair $(Dec, Conf)$ by its corresponding qualitative values. In the following example, we present two situations where the impact of changing the decision scale is clearly visible on the outcomes. Let's consider the two arguments C-Arg and D-Arg of the figure 4.15. The conclusion (C) is supported by two premises to which we assign the assessment "*opposable, for sure*" for both. Similarly, the conclusion (C') is supported by two premises $(P'_1,\ P'_2)$ with the assessments "tolerable, for sure". Supposing that all belief in rules are maximal and two different equidistant decision scale of respectively four and five items, lets calculate the confidence in the conclusions C and C'. Table 4.6 groups the result of the confidence (in quantitative and qualitative form) in the conclusions C and C' when $|Dec| = 4$ and $|Dec| = 5$. In the case of a five-item decision scale ("*rejectable*", "*opposable*", "*no decision*", "*tolerable*" and "*acceptable*"), the result for both arguments (C-Arg and D-Arg) is *no decision, for sure* suggesting that the evidence provided by the premises contains both elements supporting and refuting the conclusions, which is not the case in this situation. On the other hand, using a four-item decision scale (without "*no decision*" item) the expected outcomes are obtained, i.e. *opposable, for sure* for (C-Arg) and *tolerable, for sure* for (D-Arg). Keep in mind that C-Arg and D-Arg respectively propagate the premise with the least and the highest strength. However, the use of an odd decision scale (e.g., $|Dec| = 5$) is important in our assessment procedure for two reasons. First of all,

this central element highlights both the situation of total ignorance and the presence of conflict, as we have seen in the previous chapters, which is impossible with a even scale (e.g., $|Dec| = 4$). Secondly, this item is essential to the elicitation procedure which use ignorance on some premises to deduce beliefs on simple rules.

Moreover, through the various argument structures (safety, assurance cases, etc.) that we were able to review, we noticed that most of these cases corresponded to conjunctive argument types (C-Arg). This observation may cast doubt on the utility of having a disjunctive (D-Arg) and a hybrid (H-Arg) argument type, although these have been defined in several works [Cyra 2011, Ayoub 2013, Guiochet 2014, Wang 2018a]. We assume two reasons for this situation. The first may be related to the design of the argument structure. The argument designer proposes only minimal evidence to prove that the system is, for instance, acceptably safe and does not try to add more evidence that could improve its safety level. A second possible explanation may be related to the assessor. Indeed, it can be assumed that experts take a cautious approach in their assessment. Hence, they tend to give extreme values when evaluating an argument by rejecting it when a piece of evidence seems less credible and do not accept a conclusion until all the evidence seems more credible. The second explanation seems more likely, especially since in our argument assessment method we put the expert in extreme situations in order to deduce the (dis-)belief in the rules. The latter can indirectly influence the expert judgment.

4.5 Conclusion

In this chapter, we presented a case study on which we applied our confidence assessment procedure for GSN. A questionnaire was used to elicited both rules (modelling phase) and premises. The elicited data was transformed to belief degrees in order to compute the confidence in the top goal (conclusion), then translated back to qualitative assessment (application phase). The results of the sensitivity analysis conducted on this case were in agreement with our expectation. Thus, we can deduce that our approach is valid. However, we notice that the qualitative/quantitative transformation can lead to inconsistent results according to choice of the numerical assessment scales (i.e, decision and confidence). Especially, when approximating the degrees of belief on the conclusion $(Bel, Disb, Uncer)$ to the pair $(Dec, Conf)$. Hence there is a need to improve our elicitation model. An interesting approach, which consists in replacing this quantitative approach by a purely qualitative one, will be explored in the next chapter.

Chapter 5

An opening towards a pure qualitative approach

Contents

5.1 Introduction 83
5.2 Introduction to qualitative capacities 84
5.3 Qualitative confidence propagation models 86
5.3.1 Simple argument (S-Arg) 86
5.3.2 Conjunctive argument (C-Arg) 87
5.3.3 Disjunctive argument (D-Arg) 88
5.3.4 Hybrid argument (H-Arg) 89
5.3.5 Conflict mass 91
5.4 Qualitative elicitation models 92
5.5 Qualitative assessment application to GSN arguments . . . 93
5.6 Quantitative Vs Qualitative assessment procedure 95
5.7 Discussion 96
5.8 Conclusion 97

5.1 Introduction

In the previous chapter, we presented a quantitative method of uncertainty assessment in GSN -based argument structures. This approach uses Dempster-Shafer theory to build on top of the argument a formal model of confidence/uncertainty propagation from premises to the conclusion expressed with logical operators (conjunction, disjunction, and negation). The confidence values, used to feed the propagation model, are quantified using belief functions in the form of a 3-tuple $(Bel, Disb, Uncer)$ on rules and premises.

However this quantification of confidence is often problematic, when it relies on expert assessments. In many cases, experts supply qualitative assessments using linguistic values like *for sure, high confidence, low confidence*, etc., which are then translated into numbers on the $[0, 1]$ scale (elicitation approach, in chapter 3). This translation is somewhat arbitrary. So, a legitimate question is whether a purely qualitative approach to uncertainty, that would be a counterpart to the belief function approach, could be promising. The idea is to avoid the quantitative encoding

of qualitative estimates. It makes all the more sense as numerical degrees of belief obtained via uncertainty propagation are often translated back to the qualitative scale, so as to make the results more palatable. So it is legitimate to investigate a qualitative approach.

This chapter presents a first step in this direction. It presents theoretical background on qualitative capacities that can be viewed as a qualitative counterpart of belief functions, based on [Dubois 2022a]. Then, it presents a modified version of confidence propagation and elicitation models using qualitative capacities. Finally, a preliminary comparison of qualitative and quantitative uncertainty propagation is proposed via an example.

5.2 Introduction to qualitative capacities

A capacity (or fuzzy measure), proposed by [Zadeh 1999], is a set function $g : 2^{\Omega} \rightarrow [0, 1]$ which respect assumptions in (1.1), especially monotonicity with respect to inclusion. Hence:

$$g(A \cup B) \geq \max(g(A), g(B)) \tag{5.1}$$

$$g(A \cap B) \leq \min(g(A), g(B)) \tag{5.2}$$

When equation (5.1) is made an equality $g(A \cup B) = \max(g(A), g(B))$, then we talk about a possibility measure (named Π). In a finite setting, possibility is formally defined by a function $\pi : \Omega \rightarrow [0, 1]$ such that:

$$\Pi(A) = \max_{\omega \in A} \pi(\omega). \tag{5.3}$$

On the other hand, when equation (5.2) is forced into an equality $g(A \cap B) = \min(g(A), g(B))$, then we talk about necessity measure (named N). Similarly to belief and plausibility functions ($Pl = 1 - Disb$), necessity can be formally defined by duality as:

$$N(A) = \min_{\omega \notin A}(1 - \pi(\omega)) = 1 - \Pi(\neg A) \tag{5.4}$$

When g respects the additivity property: $g(A \cup B) = g(A) + g(B)$, where $A \cap B = \emptyset$, $g = P$ we talk about a probability measure.

A qualitative capacity (q-capacity, for short) is a function that maps the power set of Ω to a finite totally ordered symbolic (or Linguistic) set L instead of the unit interval. Namely:

$$\begin{cases} \gamma(\emptyset) = 0, \gamma(\Omega) = 1 \\ A \subseteq B \Rightarrow \gamma(A) \leq \gamma(B) \end{cases} \tag{5.5}$$

Any q-capacity can be put in a form similar to that of a belief function:

$$\gamma(A) = \max_{B \subseteq A \neq \emptyset} \rho(B), \forall A \subseteq W, \tag{5.6}$$

where ρ is called a basic possibility assignment (BΠA) [Dubois 2019] since formally,

the function $\rho : 2^W \to L$ is such that $\max_{S \subseteq W} \rho(S) = 1$ and $\rho(\emptyset) = 0$. The value $\rho(B)$ is the strength of the piece of evidence B, contrary to the numerical case γ does not induce a unique ρ. We can always assume $\rho(W) = 1$.

The linguistic set L can take different formats (Tall - Short; high level - low level; (α, δ, λ); 0 - 1, etc.). The value $\gamma(A)$ (resp. $\gamma(\neg A)$) qualifies the support in favor of (resp. against) A, i.e. belief (resp. disbelief) in A using an element in the qualitative scale L. The pair $(\gamma(A), \gamma(\neg A))$ thus describes our epistemic stance with respect to A in terms of belief and disbelief.

Figure 5.1 presents the credibility and information orderings on pairs (belief, disbelief) including extreme cases. A proposition A is at least as credible as B if $\gamma(A) \geq \gamma(B)$ and $\gamma(\neg A) \leq \gamma(\neg B)$ (solid arrows), thus ranging from certainty of falsity $(0, 1)$ up to certainty of truth $(1, 0)$. A proposition A is at least as informed as B if $\gamma(A) \geq \gamma(B)$ and $\gamma(\neg A) \geq \gamma(\neg B)$ (dotted arrows), thus ranging from ignorance ($(0, 0)$, no information) up to conflict ($(1, 1)$, full contradictory information). In this situation, the amount of evidence supporting the conclusion is equal to the one rejecting it. The set $L \times L$ is then equipped with a bilattice structure [Ciucci 2019].

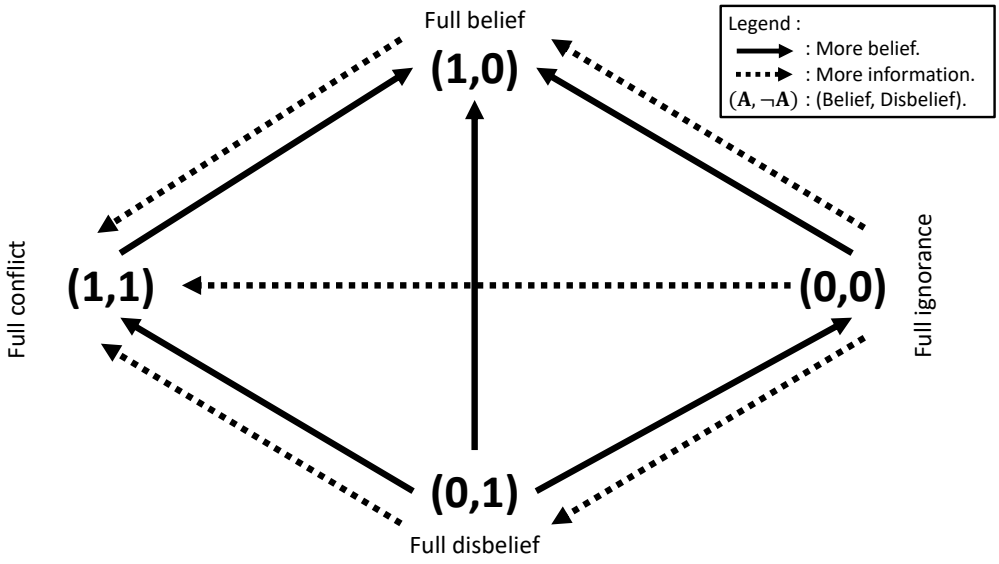


Figure 5.1: Evolution of certainty and information in pairs (belief, disbelief)

In order to combine pieces of information (evidence) represented by BΠAs (ρ_1, ρ_2, ... ρ_n), an alternative to Dempster rule of combination (equation 1.8) is given as: $\rho_\cap = \rho_1 \odot \rho_2$ such that:

$$\rho_\cap(A) = \max_{E_1 \cap E_2 = A} \min[\rho_1(E_1)), \rho_2(E_2)] \tag{5.7}$$

This is like the unnormalized Dempster rule of combination, changing product into min and sum into max.

Table 5.1: Combination of the focal sets of the premise (ρ_p) with its rules (ρ_r) for S-Arg

$\rho = \rho_p \odot \rho_r$	$\rho_r(p \equiv C)$	$\rho_r(p \Rightarrow C)$	$\rho_r(\neg p \Rightarrow \neg C)$	$\rho_r(\top)$
$\rho_p(p)$	$p \wedge C$	$p \wedge C$	p	p
$\rho_p(\neg p)$	$\neg p \wedge \neg C$	$\neg p$	$\neg p \wedge \neg C$	$\neg p$
$\rho_p(\top)$	$p \equiv C$	$p \Rightarrow C$	$\neg p \Rightarrow \neg C$	$\top$

5.3 Qualitative confidence propagation models

In this paragraph, we use the same argument types (modelled by rules, in chapter 2) to build purely qualitative propagation models. To run such models, we use an elicitation method adapted to qualitative assessments on premises and rules (presented in the next section).

5.3.1 Simple argument (S-Arg)

This argument describes the case of a conclusion (C) supported by a single premise (p). To model this argument, we associate to each rule (direct and reverse one) a simple $B\Pi A$'s (resp., $\rho_{\Rightarrow}$ and $\rho_{\Leftarrow}$), and a $B\Pi A$ on the premise space ρ_p, assigning a mass to its truth $\rho_p(p)$, its falsity $\rho_p(\neg p)$ and the tautology $\rho_p(\top) = 1$ of which their maximum is 1. Then, using a combination rule in (formula 5.7), we merge the BΠAs on rules ($\rho_r = \rho_{\Rightarrow} \odot \rho_{\Leftarrow}$) with the one on the premise (ρ_p), see table 5.1. Similarly to the quantitative formulas, $\gamma_C(C) = \rho_C(C)$, $\gamma_p(p) = \rho_p(p)$ and $\gamma_r(r) = \rho_r(r)$, $\forall r$ (conjunctive, disjunctive, direct or reverse), since we work on a two state frame of discernment $\Omega_x = \{x, \neg x\}$.

$$\begin{aligned}\gamma_C(C) &= \max_{\phi:\phi\vdash C,\phi\neq\emptyset} \rho(\phi) \\ &= \max \min[\rho_p(p), \rho_r(p \equiv C), \rho_r(p \Rightarrow C)] \\ &= \min[\rho_p(p), \rho_{\Rightarrow}(p \Rightarrow C)]\end{aligned}$$

Where: $\max \min[\rho_r(p \equiv C), \rho_r(p \Rightarrow C)] = \max[\min(\rho_{\Rightarrow}, \rho_{\Leftarrow}), \min(\rho_{\Rightarrow})] = \rho_{\Rightarrow}$.

$$\begin{aligned}\gamma_C(\neg C) &= \max_{\phi:\phi\vdash \neg C,\phi\neq\emptyset} \rho(\phi) \\ &= \max \min[\rho_p(\neg p), \rho_r(p \equiv C), \rho_r(\neg p \Rightarrow \neg C)] \\ &= \min[\rho_p(\neg p), \rho_{\Leftarrow}(\neg p \Rightarrow \neg C)]\end{aligned}$$

So, we conclude for the uncertainty propagation in simple argument:

$$\text{S-Arg} : \begin{cases} \gamma_C(C) &= \min[\gamma_p(p), \gamma_{\Rightarrow}(p \Rightarrow C)] \\ \gamma_C(\neg C) &= \min[\gamma_p(\neg p), \gamma_{\Leftarrow}(\neg p \Rightarrow \neg C)] \end{cases} \tag{5.8}$$

We can notice that the belief $\gamma_C(C)$ only depends on the direct rule and the belief of the premise, while the disbelief $\gamma_C(\neg C)$ only depends on the reverse rule and the disbelief of the premise.

5.3.2 Conjunctive argument (C-Arg)

This argument type describes the situation when two premises or more are jointly needed to support a conclusion. We formally defined its direct and reverse rules (resp.) by: $(\wedge_{i=1}^{n} p_i) \Rightarrow C$ and $\wedge_{i=1}^{n}(\neg p_i \Rightarrow \neg C)$. Following the same reasoning as for the previous argument type, we put a simple $B\Pi A$ on each rule ($\rho_{\Rightarrow}$ and $\rho^i_{\Leftarrow}$), and a function $B\Pi A$ on each premise: ρ^i_p, which assigns one mass on the truth of (p_i), its negation $(\neg p_i)$ and the tautology $(\top)$ such $\rho^i_p(\top) = 1$. Then, using the combination rule in 5.7, we deduce $\gamma_C(C)$ and $\gamma_C(\neg C)$ from the combination: $\rho_C = \rho_p \odot \rho_r$ (tables 5.2 and 5.3). Where: $\rho_p = \rho^1_p \odot \rho^2_p$ and $\rho_r = \rho_{\Rightarrow} \odot \rho^1_{\Leftarrow} \odot \rho^2_{\Leftarrow}$.

Table 5.2: Combination of masses of rules and premises for the C-Arg (part 1)

$\rho = \rho_p \odot \rho_r$	$\rho_r([p_1 \wedge p_2] \equiv C)$	$\rho_r([p_1 \wedge p_2] \Rightarrow C)$	$\rho_r(\neg[p_1 \wedge p_2] \Rightarrow \neg C)$	$\rho_r(\neg p_1 \Rightarrow \neg C)$
$\rho_p(p_1 \wedge p_2)$	$p_1 \wedge p_2 \wedge C$	$p_1 \wedge p_2 \wedge C$	-	-
$\rho_p(\neg p_1 \wedge p_2)$	$\neg p_1 \wedge p_2 \wedge \neg C$	-	$\neg p_1 \wedge p_2 \wedge \neg C$	$\neg p_1 \wedge p_2 \wedge \neg C$
$\rho_p(p_1 \wedge \neg p_2)$	$p_1 \wedge \neg p_2 \wedge \neg C$	-	$p_1 \wedge \neg p_2 \wedge \neg C$	-
$\rho_p(\neg p_1 \wedge \neg p_2)$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	-	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$
$m_p(p_1)$	-	-	-	-
$\rho_p(\neg p_1)$	$\neg p_1 \wedge \neg C$	-	$\neg p_1 \wedge \neg C$	$\neg p_1 \wedge \neg C$
$\rho_p(p_2)$	-	-	-	-
$m_p(\neg p_2)$	$\neg p_2 \wedge \neg C$	-	$\neg p_2 \wedge \neg C$	-
$\rho_p(\top)$	-	-	-	-

Table 5.3: Combination of masses of rules and premises for the C-Arg (part 2)

$\rho = \rho_p \odot \rho_r$	$\rho_r(\neg p_2 \Rightarrow \neg C)$	$\rho_r(F_1)$	$\rho_r(F_2)$	$\rho_r(\top)$
$\rho_p(p_1 \wedge p_2)$	-	$p_1 \wedge p_2 \wedge C$	$p_1 \wedge p_2 \wedge C$	-
$\rho_p(\neg p_1 \wedge p_2)$	-	$\neg p_1 \wedge p_2 \wedge \neg C$	-	-
$\rho_p(p_1 \wedge \neg p_2)$	$p_1 \wedge \neg p_2 \wedge \neg C$	-	$p_1 \wedge \neg p_2 \wedge \neg C$	-
$\rho_p(\neg p_1 \wedge \neg p_2)$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	-
$\rho_p(p_1)$	-	-	-	-
$\rho_p(\neg p_1)$	$\neg p_1 \wedge \neg C$	$\neg p_1 \wedge \neg C$	-	-
$\rho_p(p_2)$	-	-	-	-
$m_p(\neg p_2)$	$\neg p_2 \wedge \neg C$	-	$\neg p_2 \wedge \neg C$	-
$\rho_p(\top)$	-	-	-	$\top$

Where:

- $F_1 = [(p_1 \wedge p_2) \Rightarrow C] \wedge [\neg p_1 \Rightarrow \neg C]$.
- $F_2 = [(p_1 \wedge p_2) \Rightarrow C] \wedge [\neg p_2 \Rightarrow \neg C]$.

For a better visualization, we have chosen to replace by (-) the focal elements that do not trigger the conclusion (C) or its negation $(\neg C)$ in the tables 5.2 and 5.3.

From the combination of q-capacity functions in tables 5.2 and 5.3, we deduce the belief degree in the conclusion $\gamma_C(C)$:

$$\begin{aligned}\gamma_C(C) &= \max_{\phi:\phi\vdash C,\ \phi\neq\emptyset} \rho(\phi) \\ &= \min[\rho_p^1(p_1), \rho_p^2(p_2), \rho_{\Rightarrow}([p_1 \wedge p_2] \Rightarrow C)]\end{aligned}$$

From the combination of q-capacity functions in tables 5.2 and 5.3, we deduce the disbelief degree in the conclusion $\gamma_C(\neg C)$:

$$\begin{aligned}\gamma_C(\neg C) &= \max_{\phi:\phi\vdash \neg C,\ \phi\neq\emptyset} \rho(\phi) \\ &= \max[\min(\rho_p^1(\neg p_1), \rho_{\Leftarrow}^1(\neg p_1 \Rightarrow \neg C)), \min(\rho_p^2(\neg p_2), \rho_{\Leftarrow}^2(\neg p_2 \Rightarrow \neg C)]\end{aligned}$$

Generalizing the calculations above to n premises, we get the following qualitative confidence propagation formulas for a C-Arg:

$$\text{C-Arg} : \begin{cases} \gamma_C(C) &= \min\{\min_{i=1}^n \gamma_p^i(p_i), \gamma_{\Rightarrow}([\wedge_i^n p_i] \Rightarrow C)\} \\ \gamma_C(\neg C) &= \max_{i=1}^n\{\min[\gamma_p^i(\neg p_i), \gamma_{\Leftarrow}^i(\neg p_i \Rightarrow \neg C)]\} \end{cases} \tag{5.9}$$

In the formulas of the quantitative approach (formulas in Section 2.2) they use probabilistic sum ($a+b-ab$) and the product (ab) instead of max, min, highlighting the similarity between the results obtained in both models. In fact, we can better see with min-max operators that the C-Arg favors the propagation of the premise with the least strength (minimal belief, with a maximal disbelief degree).

5.3.3 Disjunctive argument (D-Arg)

In this situation, each premise can support alone the whole conclusion. Formally, the direct and reverse rules are defined as follows: $\wedge_{i=1}^n(p_i \Rightarrow C)$ and $(\wedge_{i=1}^n \neg p_i) \Rightarrow \neg C$. The calculation of $\gamma_C(C)$ and $\gamma_C(\neg C)$ is identical to the one above, swapping the two expressions ($\rho_C = \rho_p \odot \rho_r$, where $\rho_r = \rho_{\Rightarrow}^1 \odot \rho_{\Rightarrow}^2 \odot \rho_{\Leftarrow}$, tables 5.4 and 5.5).

Table 5.4: Combination of masses of rules and premises for the D-Arg (part 1)

$\rho = \rho_p \odot \rho_r$	$m_r([p_1 \vee p_2] \equiv C)$	$\rho_r([p_1 \vee p_2] \Rightarrow C)$	$\rho_r(\neg[p_1 \vee p_2] \Rightarrow \neg C)$	$\rho_r(p_1 \Rightarrow C)$
$\rho_p(p_1 \wedge p_2)$	$p_1 \wedge p_2 \wedge C$	$p_1 \wedge p_2 \wedge C$	-	$p_1 \wedge p_2 \wedge C$
$\rho_p(\neg p_1 \wedge p_2)$	$\neg p_1 \wedge p_2 \wedge C$	$\neg p_1 \wedge p_2 \wedge C$	-	-
$\rho_p(p_1 \wedge \neg p_2)$	$p_1 \wedge \neg p_2 \wedge C$	$p_1 \wedge \neg p_2 \wedge C$	-	$p_1 \wedge \neg p_2 \wedge C$
$\rho_p(\neg p_1 \wedge \neg p_2)$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	-	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	-
$\rho_p(p_1)$	$p_1 \wedge C$	$p_1 \wedge C$	-	$p_1 \wedge C$
$\rho_p(\neg p_1)$	-	-	-	-
$\rho_p(p_2)$	$p_2 \wedge C$	$p_2 \wedge C$	-	
$\rho_p(\neg p_2)$	-	-	-	-
$\rho_p(\top)$	-	-	-	-

Table 5.5: Combination of masses of rules and premises for the D-Arg (part 2)

$\rho = \rho_p \odot \rho_r$	$\rho_r(p_2 \Rightarrow C)$	$\rho_r(F_1')$	$\rho_r(F_2')$	$\rho_r(\top)$
$\rho_p(p_1 \wedge p_2)$	$p_1 \wedge p_2 \wedge C$	$p_1 \wedge p_2 \wedge C$	$p_1 \wedge p_2 \wedge C$	-
$\rho_p(\neg p_1 \wedge p_2)$	$\neg p_1 \wedge p_2 \wedge C$	-	$\neg p_1 \wedge p_2 \wedge C$	-
$\rho_p(p_1 \wedge \neg p_2)$	-	$p_1 \wedge \neg p_2 \wedge C$	-	-
$\rho_p(\neg p_1 \wedge \neg p_2)$	-	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	-
$\rho_p(p_1)$	-	$p_1 \wedge C$	-	-
$\rho_p(\neg p_1)$	-	-	-	-
$\rho_p(p_2)$	$p_2 \wedge C$	-	$p_2 \wedge C$	-
$\rho_p(\neg p_2)$	-	-	-	-
$\rho_p(\top)$	-	-	-	$\top$

Where:

- $F_1' = [\neg(p_1 \wedge p_2) \Rightarrow \neg C] \wedge [p_1 \Rightarrow C]$.
- $F_2' = [\neg(p_1 \wedge p_2) \Rightarrow \neg C] \wedge [p_2 \Rightarrow C]$.

From the combination of q-capacity functions in tables 5.4 and 5.5, we deduce the belief degree in the conclusion $\gamma_C(C)$:

$$\begin{aligned}\gamma_C(C) &= \max_{\phi:\phi\vdash C,\ \phi\neq\emptyset} \rho(\phi) \\ &= \max[\min(\rho_p^1(p_1), \rho_{\Rightarrow}^1(p_1 \Rightarrow C), \min(\rho_p^2(p_2), \rho_{\Rightarrow}^2(p_2 \Rightarrow C))]\end{aligned}$$

From the combination of q-capacity functions in tables 5.4 and 5.5, we deduce the disbelief degree in the conclusion $\gamma_C(C)$:

$$\begin{aligned}\gamma_C(\neg C) &= \max_{\phi:\phi\vdash \neg C,\ \phi\neq\emptyset} \rho(\phi) \\ &= \min[\rho_p^1(\neg p_1), \rho_p^2(\neg p_2), \rho_{\Leftarrow}([\neg p_1 \wedge \neg p_2] \Rightarrow \neg C)]\end{aligned}$$

Generalizing the calculations above to n premises, we get the following qualitative confidence propagation formulas for a D-Arg:

$$\text{D-Arg} : \begin{cases} \gamma_C(C) &= \max_{i=1}^n\{\min[\gamma_p^i(p_i), \gamma_{\Rightarrow}^i(p_i \Rightarrow C)]\} \\ \gamma_C(\neg C) &= \min\{\min_{i=1}^n \gamma_p^i(\neg p_i), \gamma_{\Leftarrow}([\wedge_i^n \neg p_i] \Rightarrow \neg C)\} \end{cases} \tag{5.10}$$

We can notice that this model, as its quantitative counterpart (formulas in 2.3), favors the propagation of the premise with the greatest strength (maximal belief and minimal disbelief degree).

5.3.4 Hybrid argument (H-Arg)

This argument type describes the situation where each premise supports the conclusion to some degree, but their conjunction does it to a larger one. Therefore, all conjunctive and elementary rules will be used in this argument type. Thus, we obtain: $\gamma_C = \rho_p^1 \odot \rho_p^2 \odot \rho_{\Rightarrow} \odot \rho_{\Rightarrow}^1 \odot \rho_{\Rightarrow}^2 \odot \rho_{\Leftarrow} \odot \rho_{\Leftarrow}^1 \odot \rho_{\Leftarrow}^2$ (table 5.6).

Table 5.6: Combination of premise (p_i) with its rules for H-Arg

$\rho = \rho_{12} \odot \rho_r$	$\rho_r(F_c)$	$\rho_r([p_1 \wedge p_2] \Rightarrow C)$	$\rho_r([\neg p_1 \wedge \neg p_2] \Rightarrow \neg C)$	$\rho_r(\top)$
$\rho_{12}(p_1 \wedge p_2 \wedge C)$	$p_1 \wedge p_2 \wedge C$	$p_1 \wedge p_2 \wedge C$	$p_1 \wedge p_2 \wedge C$	$p_1 \wedge p_2 \wedge C$
$\rho_{12}(\neg p_1 \wedge p_2 \wedge C)$	$\neg p_1 \wedge p_2 \wedge C$	$\neg p_1 \wedge p_2 \wedge C$	$\neg p_1 \wedge p_2 \wedge C$	$\neg p_1 \wedge p_2 \wedge C$
$\rho_{12}(p_1 \wedge \neg p_2 \wedge C)$	$p_1 \wedge \neg p_2 \wedge C$	$p_1 \wedge \neg p_2 \wedge C$	$p_1 \wedge \neg p_2 \wedge C$	$p_1 \wedge \neg p_2 \wedge C$
$\rho_{12}(p_1 \wedge C)$	$p_1 \wedge C$	$p_1 \wedge C$	$p_1 \wedge C$	$p_1 \wedge C$
$\rho_{12}(p_2 \wedge C)$	$p_2 \wedge C$	$p_2 \wedge C$	$p_2 \wedge C$	$p_2 \wedge C$
$\rho_{12}(\neg p_1 \wedge \neg p_2 \wedge \neg C)$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$
$\rho_{12}(\neg p_1 \wedge p_2 \wedge \neg C)$	$\neg p_1 \wedge p_2 \wedge \neg C$	$\neg p_1 \wedge p_2 \wedge \neg C$	$\neg p_1 \wedge p_2 \wedge \neg C$	$\neg p_1 \wedge p_2 \wedge \neg C$
$\rho_{12}(p_1 \wedge \neg p_2 \wedge \neg C)$	$p_1 \wedge \neg p_2 \wedge \neg C$	$p_1 \wedge \neg p_2 \wedge \neg C$	$p_1 \wedge \neg p_2 \wedge \neg C$	$p_1 \wedge \neg p_2 \wedge \neg C$
$\rho_{12}(\neg p_1 \wedge \neg C)$	$\neg p_1 \wedge \neg C$	$\neg p_1 \wedge \neg C$	$\neg p_1 \wedge \neg C$	$\neg p_1 \wedge \neg C$
$\rho_{12}(\neg p_2 \wedge \neg C)$	$\neg p_2 \wedge \neg C$	$\neg p_2 \wedge \neg C$	$\neg p_2 \wedge \neg C$	$\neg p_2 \wedge \neg C$
$\rho_{12}(p_1 \wedge p_2)$	$p_1 \wedge p_2 \wedge C$	$p_1 \wedge p_2 \wedge C$	-	-
$\rho_{12}(\neg p_1 \wedge \neg p_2)$	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	-	$\neg p_1 \wedge \neg p_2 \wedge \neg C$	-
...	...	...	...	...
$\rho_{12}(\emptyset)$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
$\rho_{12}(\top)$	F_c	$[p_1 \wedge p_2] \Rightarrow C$	$[\neg p_1 \wedge \neg p_2] \Rightarrow \neg C$	$\top$

Where:

- $\rho_{12} = (\rho_p^1 \odot \rho_{\Rightarrow}^1 \odot \rho_{\Leftarrow}^1) \odot (\rho_p^2 \odot \rho_{\Rightarrow}^2 \odot \rho_{\Leftarrow}^2)$
- $\rho_r = \rho_{\Rightarrow} \odot \rho_{\Leftarrow}$
- $F_c = ([p_1 \wedge p_2] \Rightarrow C) \wedge ([\neg p_1 \wedge \neg p_2] \Rightarrow \neg C)$.

From the combination of q-capacity functions in table 5.6, we deduce the belief degree in the conclusion $\gamma_C(C)$:

$$\begin{aligned}\gamma_C(C) &= \max_{\phi:\phi\vdash C,\ \phi\neq\emptyset} \rho(\phi) \\ &= \max\{\min[\rho_p^1(p_1), \rho_p^2(p_2), \rho_{\Rightarrow}([p_1 \wedge p_2] \Rightarrow C)], \\ &\qquad \min[(\rho_p^1(p_1), \rho_{\Rightarrow}^1(p_1 \Rightarrow C)), \min(\rho_p^2(p_2), \rho_{\Rightarrow}^2(p_2 \Rightarrow C)]\}\end{aligned}$$

From the combination of q-capacity functions in table 5.6, we deduce the disbelief degree in the conclusion $\gamma_C(C)$:

$$\begin{aligned}\gamma_C(\neg C) &= \max_{\phi:\phi\vdash \neg C,\ \phi\neq\emptyset} \rho(\phi) \\ &= \max\{\min[\rho_p^1(\neg p_1), \rho_p^2(\neg p_2), \rho_{\Leftarrow}([\neg p_1 \wedge \neg p_2] \Rightarrow \neg C)], \\ &\qquad \min[(\rho_p^1(\neg p_1), \rho_{\Leftarrow}^1(\neg p_1 \Rightarrow \neg C)), \min(\rho_p^2(\neg p_2), \rho_{\Leftarrow}^2(\neg p_2 \Rightarrow \neg C)]\}\end{aligned}$$

Generalizing the calculations above to n premises, we get the following qualitative confidence propagation formulas for a H-Arg:

$$\text{H-Arg}: \begin{cases} \gamma_C(C) = & \max\{\min[\min_{i=1}^n \gamma_p^i(p_i), \gamma_{\Rightarrow}([\wedge_{i=1}^n p_i] \Rightarrow C)], \\ & \max_{i=1}^n(\min[\gamma_p^i(p_i), \gamma_{\Rightarrow}^i(p_i \Rightarrow C)]\} \\ \gamma_C(\neg C) = & \max\{\min[\min_{i=1}^n \gamma_p^i(\neg p_i), \gamma_{\Leftarrow}([\wedge_{i=1}^n \neg p_i] \Rightarrow \neg C)], \\ & \max_{i=1}^n \min[\gamma_p^i(\neg p_i), \gamma_{\Leftarrow}^i(\neg p_i \Rightarrow \neg C)]\} \end{cases} \tag{5.11}$$

We can notice, as expected (analogy to quantitative formulas), that formulas of H-Arg (5.11), presents a combination between C-Arg formulas (5.9), and D-Arg (5.10). Assuming a maximal belief (= 1) (resp. disbelief) on premises, it is enough that the simple direct rules take a null value (resp. the reversed conjunctive one) to get the conjunctive argument type. And conversely, to get the disjunctive argument type, put null values on direct conjunctive and simple reversed rules. The S-Arg, represents a special case when only one premise is available ($n = 1$). In the following, only the H-Arg will be used since it covers the four types.

5.3.5 Conflict mass

The C-Arg and D-Arg argument type are conflict-free. They respectively propagate the premise with lowest and highest assessments. However, it is not the same for H-Arg. Indeed, merging $B\Pi A$'s ρ_p^i (on p_i, $\neg p_i$ and $\top$), $\rho_{\Rightarrow}^i$, $\rho_{\Leftarrow}^i, i = 1, \ldots n$, as in its quantitative counterpart, the BΠA pertaining to the conclusion C obtained from this fusion may assign a mass to the contradiction. Conflict always appears when four items are merged of the form: p_i and $p_i \Rightarrow C$ with $\neg p_j$ and $\neg p_j \Rightarrow \neg C, j \neq i$,

whose conjunction is a contradiction $\emptyset$ with mass:

$$\begin{aligned}\rho_C^{ij}(\emptyset) &= \min[\rho_C^i(p_i \wedge C), \rho_C^j(\neg p_j \wedge \neg C)] \\ &= \min[\rho_p^i(p_i), \rho_{\Rightarrow}^i(p_i \Rightarrow C), \rho_p^j(\neg p_j), \rho_{\Rightarrow}^j(\neg p_j \Rightarrow \neg C)]]\end{aligned}$$

The final capacity on contradiction for n premises takes the form $\rho_C(\emptyset) = \max_{i \neq j}(\rho_C^{ij}(\emptyset))$. Besides, this capacity on contradiction does not affect the final results of belief and disbelief since $\gamma_C(C)$ and $\gamma_C(\neg C)$ are greater than $\gamma_C(\emptyset)$. For instance, for $n = 2$: $\rho_C(\emptyset) = \max(\rho_C^{12}(\emptyset), \rho_C^{21}(\emptyset))$. Using (5.6) we get: $\gamma_C(C) = \max[\rho_C(p_1 \wedge C), \rho_C(p_2 \wedge C)] \geq \rho_C(\emptyset)$ and $\gamma_C(\neg C) = \max[\rho_C(\neg p_1 \wedge \neg C), \rho_C(\neg p_2 \wedge \neg C)] \geq \rho_C(\emptyset)$. Notice that: $\rho_C(p_i \wedge C) = \min(\rho_p^i(p_i), \rho_{\Rightarrow}^i(p_i \Rightarrow C)$ and $\rho_C(\neg p_i \wedge \neg C) = \min(\rho_p^i(\neg p_i), \rho_{\Leftarrow}^i(\neg p_i \Rightarrow \neg C)$.

5.4 Qualitative elicitation models

In order to elicit qualitative capacities, we reuse a modified version of the quantitative one proposed in chapter 3. Thus, the same types of information will be collected to assess the GSN pattern:

- The decision index denoted by $Dec(A)$, describes which side the assessor leans towards, i.e., acceptance or rejection of A. It is associated with a bipolar scale $D = \{0_D = d_{-n}, d_{n-1}, \ldots, d_0 = e, d_1, \ldots d_n = 1_D\}$ with $2n + 1$ values, the bottom of which (0_D) expressing rejection, the top (1_D) acceptance, and the midpoint (e) a neutral position. Here we assume $n = 2$.

- The confidence index denoted by $Conf(A)$ reflects the amount of information an assessor possesses to support the decision. It uses a positive uni-polar scale with $n + 1$ values $C = \{0_C = c_0, c_1, \ldots, c_n = 1_C\}$ (the top 1_C expresses full confidence, the bottom 0_C is neutral- no information). For $n = 2$: lack of confidence, moderate confidence and full confidence.

To make decision and confidence scales compatible with the transformation formulas presented below, we consider that:

(i) The bipolar scale D is equipped with an order-reversing map ν_D such that $\nu_D(d_{-i}) = d_i$. Especially we have that $\nu_D(Dec(A)) = Dec(\neg A)$.

(ii) The unipolar scale C is isomorphic to the positive part of D, and is equipped with an order-reversing map ν such that: $\nu(c_i) = c_{n-i}$.

In order to switch from a pair $(Dec(A), Conf(A))$ to a pair of capacity values $(\rho(A), \rho(\neg A))$, we define transformation formulas that map $D \times C$ to the belief-disbelief scale $L \times L$ containing pairs $(\gamma(A), \gamma(\neg A))$. The scale L has the same number of elements as C (i.e., 3 here). The mapping $f : D \times C \to L \times L$: must satisfy some conditions [Dubois 2022b]:

- If the expert declares lack of confidence, the result is $f(Dec(A), 0) = (0, 0)$, whatever the trend expressed on the decision scale.

Table 5.7: Values from (Dec, $Conf$) to (Bel,$Disb$) pairs on premises

$Conf$ \ Dec	d_{-2} (Rej)	d_{-1} (Opp)	$d_0(ND)$	d_1 (Tol)	d_2 (Acc)
c_0 (Lack of conf.)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)
c_1 (Moderate conf.)	(0,λ)	(λ,λ)	(λ,λ)	(λ,λ)	(λ,0)
c_2 (For sure)	(0,1)	(λ,1)	(1,1)	(1,λ)	(1,0)

- If the expert is fully confident, then $f(1,1) = (\gamma(A), \gamma(\neg A)) = (1,0)$, $f(0,1) = (0,1)$, $f(e,1) = (1,1)$. Indeed, for the latter, there is a total conflict: the expert is maximally informed, and cannot decide between A and its negation.
- $\max(\gamma(A), \gamma(\neg A)) = Conf(A)$: the belief in A or its negation cannot be stronger than the confidence.
- if $Dec(A)$ is the midpoint of D, then $\gamma(A) = \gamma(\neg A)(= Conf(A))$ (no reason to take side).
- if $Dec(A)$ is less than the midpoint of D, then $\gamma(A) < \gamma(\neg A) = Conf(A)$, and the smaller $D(A)$, the smaller $\gamma(A)$.
- if $Dec(A)$ is greater than the midpoint of D, then $\gamma(A) = Conf(A) > \gamma(\neg A)$, and the greater $D(A)$, the smaller $\gamma(\neg A)$.

These conditions lead to propose the following translation formulas :

if $Dec(A) < e,\ \gamma(A) = \min[\nu_C(Dec(\neg A)), Conf(A)]$ and $\gamma(\neg A) = Conf(A)$
if $Dec(A) > e,\ \gamma(A) = Conf(A)$ and $\gamma(\neg A) = \min[\nu_C(Dec(A)), Conf(A)]$
if $Dec(A) = Dec(\neg A) = e, \quad \gamma(A) = \gamma(\neg A) = Conf(A)$

In table 3.1, we grouped all possible (Dec, $Conf$) pairs on premises with their appropriate ($\gamma(A)$, $\gamma(\neg A)$) counterparts, using the formulas above. We can notice an anti-symmetry between belief and disbelief degrees with respect to the central column (d_0: No clear decision). We also notice that when no information is available (c_0: Lack of confidence), no matter what choice is made, the degrees of belief and disbelief take the minimal value. On the other hand, in the case of a fully informed expert (c_2: For sure) the decision value varies from rejection to acceptance and is reflected by the pair $(\gamma(A), \gamma(\neg A))$. We can see that the values in the table respect the conditions imposed above.

5.5 Qualitative assessment application to GSN arguments

Similarly to the quantitative assessment procedure, this procedure is also structured in two phases as presented in figure 5.2. A *modeling phase* where collected expert

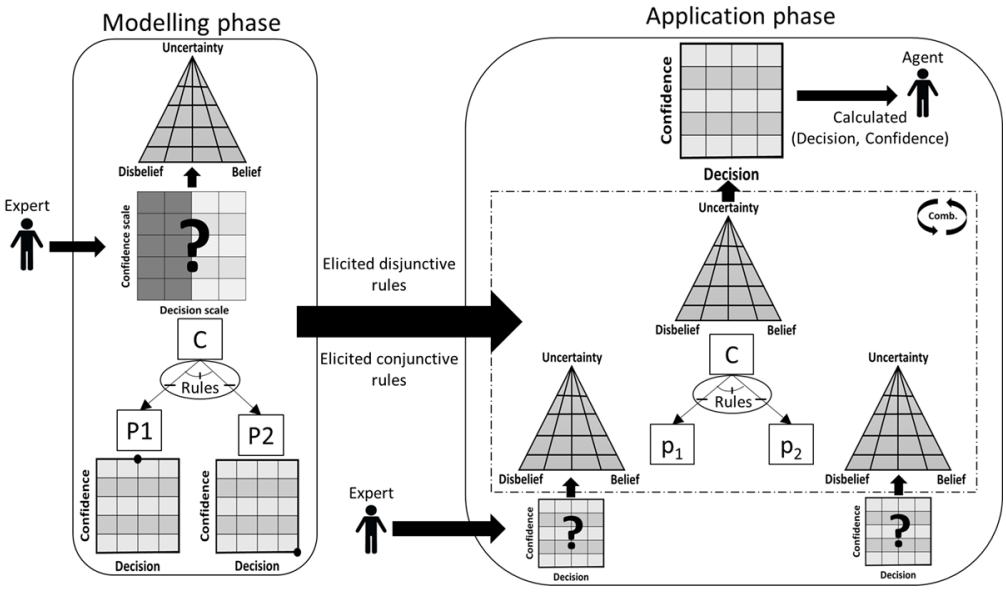


Figure 5.2: Schema of the assessment framework for safety argument

opinions about rules expressed with qualitative scores $(Dec, Conf)$, are translated to qualitative belief and disbelief values (λ_i). We benefit from an observation made on the quantitative model (formulas 2.4). Namely, we notice that under some assumptions on the premises, the value of the conclusion is the value of the rule. For instance, assuming full support (resp. positive or negative) on all premises gives the value of the conjunctive rule (resp. direct and reversed): $\gamma_C(C) = \gamma_{\Rightarrow}([\wedge_{i=1}^{n} p_i] \Rightarrow C)$ or $\gamma_C(\neg C) = \gamma_{\Leftarrow}([\wedge_{i=1}^{n} \neg p_i] \Rightarrow \neg C)$. On the other hand, assuming a total support (resp. positive or negative) on one premise (p_i) and total ignorance on the others gives the value of the appropriate disjunctive rule: $\gamma_C(C) = \gamma_{\Rightarrow}^{i}(p_i \Rightarrow C)$ or $\gamma_C(\neg C) = \gamma_{\Leftarrow}^{i}(\neg p_i \Rightarrow \neg C)$.

So, we will use the same Table 5.7 to transform the assessment on rules. However, to avoid the negation of rules, the assessor can again only choose between the positive decision levels (from *"no decision"* to *"acceptable"*) for direct rules; only negative decisions (from *"rejectable"* to *"no decision"*) for the reversed ones. Indeed, we have seen that rules can only infer uncertainty on one side of the decision scale.

Next, comes the *application phase* where assessments about premises are collected according to the same elicitation method on rules. This time the expert is free to choose any value of the evaluation matrix. Once the beliefs about the rules and premises are collected and transformed into an appropriate form, we calculate the degree of belief and disbelief of the conclusion.

Table 5.8: Values of the belief degrees on rules

Goal (G_i)	Belief degree on rules in quantitative (qualitative)
G	$Bel_{\Rightarrow}([G_1 \wedge G_2] \Rightarrow G) = Bel_{\Leftarrow}([\neg G_1 \wedge \neg G_2] \Rightarrow \neg G) = 1$ (1) $Bel_{\Rightarrow}(G_1 \Rightarrow G) = Bel_{\Rightarrow}(G_2 \Rightarrow G) = 0.75$ (λ) $Bel_{\Leftarrow}(\neg G_1 \Rightarrow \neg G) = Bel_{\Leftarrow}(\neg G_2 \Rightarrow \neg G) = 0.75$ (λ)
G_1	$Bel_{\Rightarrow}([P_1 \wedge P_2] \Rightarrow G_1) = 1$ (1) $Bel_{\Leftarrow}(\neg P_1 \Rightarrow \neg G_1) = Bel_{\Leftarrow}(\neg P_2 \Rightarrow \neg G_1) = 1$ (1)
G_2	$Bel_{\Rightarrow}(P_3 \Rightarrow G_2) = Bel_{\Rightarrow}(P_4 \Rightarrow G_2) = 1$ (1) $Bel_{\Leftarrow}([\neg P_3 \wedge \neg P_4] \Rightarrow \neg G_2) = 1$ (1)

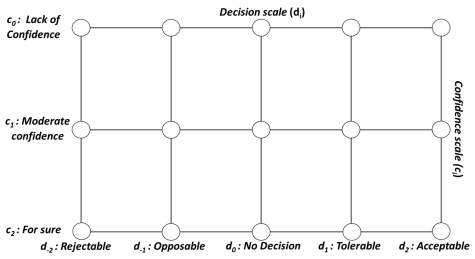


Figure 5.3: Evaluation matrix

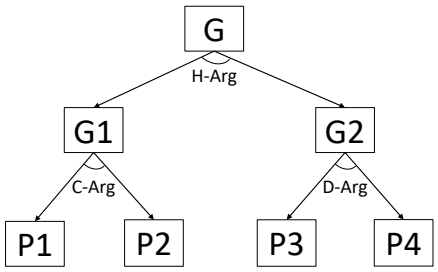


Figure 5.4: GSN artificial example

5.6 Quantitative Vs Qualitative assessment procedure

On an artificial example (Figure 5.4) that displays three argument types (C-Arg, D-Arg and H-Arg), we apply our approach in order to see how each type affects the propagation of uncertainty from premises to the overall goal (conclusion). We also apply the quantitative approach presented in chapter 4 on this artificial case study. To compare results from both approaches, we will use the same decision and confidence scales presented in figure 5.3.

The example in Figure 5.4, presents a top-goal (G) supported by two sub-goals (G1) and (G2) through a hybrid argument type (H-Arg). Each one of them is also supported, respectively, by two premises. Goal (G1) is supported by the premises (P1) and (P2) related by a conjunctive argument type (C-Arg). On the other hand, goal (G2) is supported by the premises (P3) and (P4) related by a disjunctive argument type (D-Arg). For simplicity, we set all masses on rules of C-Arg, D-Arg, and the conjunctive ones of H-Arg to their maximal values ("acceptable for sure"), see table 5.8. While we set disjunctive rule masses of H-Arg to "tolerable, for sure". Then, we use four settings with different premise values and compute the confidence in the top goal.

The procedure described in section 5.4 was used to transform the assessments into q-capacities for the qualitative case. The formulas in 3.2 and 3.1 were used to get belief and disbelief for the quantitative one.

In general, we can see from Table 5.9 that both approaches give close results which fit well with our expectations. The only difference is in the confidence values.

Table 5.9: Pairs (decision, confidence) according to both qualitative (Qual.) and quantitative (Quant.) methods for the example (see Fig. 5.3 for the meaning of symbols)

	Meth.	1^{st}	2^{nd}	3^{th}	4^{th}
P_1	-	(*Opp*,c_2)	(*Tol*,c_2)	(*Tol*,c_2)	(*Opp*,c_2)
P_2	-	(*Tol*,c_2)	(*Tol*,c_2)	(*Tol*,c_2)	(*Opp*,c_2)
P_3	-	(*Tol*,c_2)	(*Tol*,c_2)	(*Tol*;c_2)	(*Opp*;c_2)
P_4	-	(*Tol*,c_2)	(*Opp*,c_2)	(*Tol*,c_2)	(*Opp*,c_2)
G	Quant.	(*ND*;c_1)	(*Tol*;c_1)	(*Tol*;c_1)	(*Opp*;c_1)
	Qual.	$(\lambda, \lambda) \equiv$ (*ND*;c_1)	(*Tol*;c_2)	(*Tol*;c_2)	(*Opp*;c_2)

We can say that, in this case the qualitative approach gives results with higher levels of confidence (more optimistic) than the quantitative one.

Indeed, we notice that the first test case gives a "*no decision*" for the quantitative case. In the qualitative one, the formulas gives the following result: $\gamma_C(C) = \gamma_C(\neg C) = \lambda$. This result corresponds to three assessment values (see table 4.4): "opposable", "no decision" and "tolerable" with moderate confidence. Since we end up with two opposite judgments in the H-Arg (conflict situation, for both qualitative and quantitative cases) due to C-Arg that propagates the premise with least strength (opposable) to G_1 (G_2: tolerable), the decision value of the conclusion G will take also "*no decision*". On the contrary, in the 2^{nd} column, we get a "*tolerable*" decision, because the D-Arg favors the propagation of the premise with the greatest weight (tolerable) to G_2 (G_1: tolerable). In the 3^{th} and 4^{th} columns we can notice, as expected, that the top goal keeps the same decision as premises respectively: "*tolerable*" and "*opposable*".

The difference in the degree of confidence between qualitative and quantitative approaches is due to the nature of the operations used. For example, the C-Arg favors the propagation of the weakest premise (weaker belief and stronger disbelief). In the quantitative setting, we use the product and the probabilistic sum. And in the qualitative case, we use min and max, which does not model attenuation or reinforcement effects in case of independent pieces of information. This is one limitation of the qualitative approach.

5.7 Discussion

It is reasonable to assume that the more evidence we get from an argument structure, the more confidence we have in the system it supports. Hence the interest to develop an hybrid argument type.. This argument increases the confidence in the supported system to the acceptance or rejection of the conclusion according to the evidence provided (resp. for or against). It also decreases the confidence in case of lack of evidence or contradictory evidence. As we have seen from the definition of the H-Arg, this argument takes into consideration both effects of conjunction and disjunction of premises on the conclusion. However, in the end it delivers the confidence induced by the most dominant argument type C-Arg or D-Arg. This ob-

servation is more visible in the qualitative case where the "max" operator is applied between conjunctive and disjunctive formulas (see, equations 5.11). In the quantitative case, the confidence induced by the C-Arg and D-Arg are added together (see, equations 2.4).

To summarize, we can state that the main asset of this qualitative approach is that there is no need to transform the expert assessments, on the argument structure, into numerical values in order to propagate the uncertainty that may pervade it. The choice of a linear scale proposed in the quantitative approach (used for simplicity) may not be suitable for all types of systems. For instance in nuclear safety applications, it is common to use a logarithmic scale for severity or likelihood measures. A legitimate question that may arise in this case is whether the elicitation formulas remain appropriate for all types of scales? On the other hand, when the elicitation procedure relies on a single expert to assess the argument, it is safer to assume that the sources of information are not independent, hence the advantage of the qualitative method which, unlike the DST-based method, does not assume the independence of the sources.

In contrast, we can say that the main weakness of the qualitative approach is its lack of accuracy. Indeed, due to the limitation related to the scale choice ($|Conf| = n + 1$ and $|Dec| = 2n + 1$)[1], we end up with same belief and disbelief value for different assessments. Like in the example above (for $n = 2$, see table 5.7) the value (λ, λ) corresponds to three different assessments "*opposable*", "*no decision*" and "*tolerable*" with "*moderate confidence*", which can be quite confusing. In simple cases, one can deduce the correct pair $(Dec, Conf)$ on the conclusion by looking at the premises that led to the conclusion (e.g., first column in table 5.9). Yet, in complex cases with more than two premises, it may be hard to deduce the correct conclusion. With the quantitative approach such a problem does not exist. However, in order to give the final assessment on the conclusion, we will need to approximate the qualitative value of the pair $(Dec, Conf)$, which as we seen can also lead to unsatisfying results (based on the choice of scale).

Finally, the last point concerns the effects of reinforcement due to use of probabilistic sum $(a + b - ab)$ and attenuation due to the product $(a \cdot b)$ encountered in the quantitative approach. According to the results obtained from the example above, despite the few limitations that were raised, we can assume that the qualitative approach seems more promising than the quantitative one. Indeed, even if we generally get the same decision between these two approaches, maintaining a relatively high level of confidence in the qualitative case tips the balance in its favor. However, several tests on real cases will be necessary to assert this hypothesis.

5.8 Conclusion

In this chapter, we proposed a purely qualitative method of confidence/uncertainty assessment in GSN based argument structure. This procedure includes two mod-

[1]Note that beyond $n = 3$, it becomes quite challenging for the expert to make a choice.

els. A first model of confidence propagation that defines qualitative propagation formulas based on logic and the qualitative capacity theory to quantify and propagate confidence accordingly. Then a second model that transforms the collected assessments about the argument structure into belief and disbelief values used as inputs for the propagation model. A preliminary experimentation was conducted on a generic example to validate and compare this approach with its quantitative counterpart. The preliminary results showed similarities between the two, which allowed us to validate the purely qualitative approach. However, more experiments on real case studies are needed to determine the best method.

Conclusion

In this thesis, we studied the issue of assessing uncertainty (or confidence) in argument structures modeled using "*Goal Structuring Notation*" (GSN), in particular the one focusing on the safety of systems known as "*safety cases*". Thus, based on previous works from the literature, we proposed a method that exploits assessments, elicited from experts, of the evidence supporting the argument, and transforms them into degrees of belief and disbelief that are propagated to the top claim (to be supported) in order to validate it.

We conclude this thesis by recalling the different steps that we followed to elaborate such a method, while pointing out the novelties that it proposes in comparison with the existing approaches. We also discuss our perspectives for future work.

Summary

Argument structures are proposed to justify some high-level properties (e.g., safety, security, etc.) of critical system. Some safety standards (e.g., the railway standard EN50129) even request the use of a safety argument structure, also known as a "safety case", to demonstrate compliance with safety requirements. However, the evidence used to justify these high-level properties can be open to doubt, notably when it is provided by unsuitable professionals or conducted under unsuitable conditions. On the other hand, even if we assume that the evidence provided is trustful, one can always question its relevance for supporting the claim it supports (e.g., safety). Hence there is a need to build a confidence propagation model in parallel to the argument structure in order to assess the level of confidence one can attribute to it. An argument with a high level of uncertainty can hinder effective decision-making.

In the first chapter, we introduced some theoretical background related to this thesis. We first defined argument structures and pointed out the benefits of using graphical notation like GSN to model them. Then, we presented some uncertainty theories with a focus on Dempster-Shafer Theory (DST) for its effective handling of uncertainties due to lack of information often encountered in argument structures. We also explored the state of the art of approaches that apply DST to quantify and propagate uncertainty, particularly in argument structures. We conclude that DST is a good candidate to model and propagate uncertainty in GSN. However, we also explain why the few current approaches from the literature should be improved

In the second chapter, we identified the two main sources of uncertainty in GSN-based arguments. The first quantifies confidence in premises, while the second quantify the support relation between those premises and the conclusion. We identify three types of support relation, named argument types: conjunctive (C-Arg), disjunctive (D-Arg) and hybrid argument type (H-Arg). We used logical expression (logical-AND, logical-OR and implication) to formally define them. We

named this formal definition of argument types "rules". Then we use belief functions and Dempster rule of combination to build confidence propagation formulas. A sensitivity analysis was conducted on these argument types to study their behaviours.

In the third chapter, we presented an expert elicitation approach to collect and transform assessments on the argument structure into belief degrees. These values can be used by the propagation model to compute the overall confidence in the conclusion. These assessments are collected in the form of a decision (from acceptance of the claim to its rejection), and the confidence that can be associated to it. The numerical values associated to each of these qualitative items (decision, confidence) are then transformed into 3-tuples: belief, disbelief and uncertainty. The model also allows the inverse transformation to recover the assessment of the conclusion. Finally, we compared some transformation formulas from belief functions to probabilities to justify the use of our elicitation model.

In the fourth chapter we proposed a confidence assessment process for GSN-based argument structures. This process uses the elicitation approach by launching a questionnaire (for experts) to provide degrees of belief for rules. These elicited values are used in the propagation model. This model can then be used for any values of belief in the premises (evidence), also collected from experts, to compute the overall confidence in the conclusion. We also presented an application on a case study.

In the fifth chapter, we have presented a new, purely qualitative method of assessing confidence in argument structures. It uses a qualitative counterpart of belief functions, known as qualitative capacities theory, to quantify and propagate confidence in GSN. Thus, we proposed new propagation formulas, adapted to this new tool, for the three argument types we defined in the second chapter. This approach uses min-max operators to propagate qualitative belief values to the conclusion. Since it is no longer necessary to transform the assessments of the argument structure into numerical values, we use the same questionnaire to collect these assessments. We closed this chapter with a comparison between the quantitative and qualitative methods.

Main contributions

The main contributions of our research work are listed as follows:

1. Formal definition of argument types and new rigorously established confidence propagation formulas using DST:

   We formally define argument types using logical expressions. Unlike previous works, in particular those Wang et al. [Wang 2019], we chose to use the logical implication to express the support relation between premises and the conclusion. This choice allowed us to assess the belief side of the argument alone, generally to validate a system in order to integrate it into a larger

system, to obtain certification or authorization to launch production, etc. On the other hand, it enables to assess the disbelief side alone for an internal evaluation to identify, for example, weaknesses in the adopted safety strategy or potential directions for improvement. When both types of information inducing belief and disbelief are available, which is not always the case, the system can be assessed from both sides using two implications.

We also use a single confidence propagation formula to compute the confidence in the conclusion for the three types of arguments we have defined. Thus, we no longer need to use a method to identify the type of argument in order to choose the appropriate formula to apply. Keep in mind that the conjunctive (C-Arg) and the disjunctive (D-Arg) argument type represent limit cases of the hybrid (H-Arg). According to the values taken by the degrees of belief of the rules (direct and inverse), the type is easily identifiable. This allows for easy implementation and fast execution of this model for possible future applications.

2. Consideration of conflict situations in the calculation of confidence in the conclusion:

   Unlike previous works, our propagation model takes in consideration conflict situations where an expert gives two opposite assessments of at least two different premises. This is characterized by the appearance of an empty intersection with a strictly positive mass in the calculation.

3. A new model of elicitation:

   We formally define decision as a "*Pignistic* transform" and add a neutral position to its scale ("*ND*", where $Dec = 1/2$) to address the weaknesses of the previous models which give an arbitrary result in case of full ignorance ($Dec = 1/2, Conf = 0$) and did not take into consideration the situation of full conflict ($Dec = 1/2, Conf = 1$).

   We also proposed a questionnaire to collect beliefs about the rules in addition to those about the premises, to replace the approach of Wang et al. [Wang 2018a] which can lead to masses outside the unit interval $[0, 1]$.

4. A new purely qualitative confidence assessment in GSN based argument structures:

   We proposed a new purely quantitative confidence assessment method which propagates qualitative belief and disbelief from the premises to the conclusion. This approach avoids the need for transforming expert assessments from natural language (e.g., acceptable for sure, opposable with high confidence, etc.) into numerical values which seems arbitrary and can also be seen as a source of uncertainty. Moreover, the numerical degrees of belief obtained are often brought back to natural language, so as to be more palatable, which can lead to additional errors.

Limitations and Future work

The use of DST to quantify and propagate confidence in graphical models of argument structure such as GSN is not widespread and is relatively new compared to Bayesian approaches. However, the main problem to overcome (for both Bayesian and DST bases approaches) lies in the elicitation procedure. Most of the works we encountered stop after giving the propagation formulas, which makes them inapplicable in real cases. Nevertheless, other open questions still need to be investigated and developed. In this subsection we point out these issues to a possible improvements.

- Argument types formal definition:

  The hybrid argument type that we present as a general type to calculate the degrees of belief and disbelief of different argument patterns does not take into account all the possible interactions between the premises and the conclusion. For instance in the case of a conclusion (C) supported by three premises (p_1, p_2 and p_3), we only consider rules (belief side): $p_1 \Rightarrow C$, $p_2 \Rightarrow C$, $p_3 \Rightarrow C$ and $(p_1 \wedge p_2 \wedge p_3) \Rightarrow C$; while we ignore the other possibilities: $(p_1 \wedge p_2) \Rightarrow C$, $(p_1 \wedge p_3) \Rightarrow C$ and $(p_2 \wedge p_3) \Rightarrow C$. This choice was imposed because of the complexity of obtaining the corresponding masses. Thus, more experiments need to be run to find out if this compromise between a rigorous modeling of argument types and the complexity of the elicitation procedure was effective. Hence, we plan to improve the hybrid argument type to express more faithfully the relationship between premises and conclusion; and compare it to the formulas that we proposed.

- Elicitation procedure:

  The elicitation model for both the quantitative and qualitative approach need improvements. On one hand, the choice of a linear scale for decision and confidence was only taken for simplicity purposes since their is no apparent reasons to choose a specific type of scale. It is important to develop specific scales calibrated to real assurance cases and verify that the results given by the elicitation model are still correct. On the other hand, it seems that the way in which the questions are asked in the questionnaire we proposed encourages the experts to give extreme assessments. This leads to extreme types of arguments (often C-Arg). Thus, it is also necessary to develop the questionnaire by improving, for instance, the way in which the questions are asked.

- Further validation:

  Most case studies we encountered only involved the conjunctive argument type, proposed in this manuscript. We suggested some hypotheses that might explain this situation. However, for a robust validation of our approach, more experiments on general cases must be conducted for both quantitative and qualitative assessment methods.

Conducting more experiments would also be useful to compare the quantitative assessment method and the qualitative one, since the latter was only a preliminary step to know if the two models had the same behavior, but it is not enough to determine which one gives better results. Hence, we plan to run more experiments to eventually to determine the best one.

- Extension to other graphical formalisms

  The approach we propose only takes into account the components existing in the GSN formalism (*goal, strategy, solutions, context,* etc). Thus, it is not always easy to switch from one notation to another. By integrating other components our approach will become more generic. An interesting proposal is to take into account exceptions to rules (represented by the *strategy* component), by considering mass on the negation of rules we defined (i.e., $Bel_{\Rightarrow}(\neg[p \Rightarrow C]) = Bel_{\Rightarrow}([p \wedge \neg q]) = Disb_{\Rightarrow}(p \Rightarrow C)$). The component representing these exceptions is known as *defeater*.

  Hence, we plan to extend our approach to other types of notations like CAE or SACM. We also plan to improve the elicitation approach accordingly to assess the additional information requested by these notations.

Appendix A

Questionnaire for expert elicitation

Context

The objective of this assessment is to quantify the confidence in an assurance case or a safety case as it was studied at LAAS [Idmessaoud 2020, Idmessaoud 2021a, Idmessaoud 2021b, Idmessaoud 2022]. In this questionnaire, we ask an expert to provide her/his judgment regarding some specific situations. Her/his answers will be then integrated in a quantitative framework which is not presented here.

How to play

In order to assess confidence in an assurance case, an assessor needs to evaluate all goal that leads to the overall claim. We adopt an evaluation matrix (figure A.1) to assess each goal by two criteria:

- *Decision*: In a scale of 5 equidistant items, it describes which side the expert leans towards, from the rejection of the claimed goal to its acceptance.
- *Confidence*: It reflects the amount of information an assessor possesses that can justify her/his decision. There are 6 equidistant levels of the confidence scale. From “Lack of confidence" to “For sure".

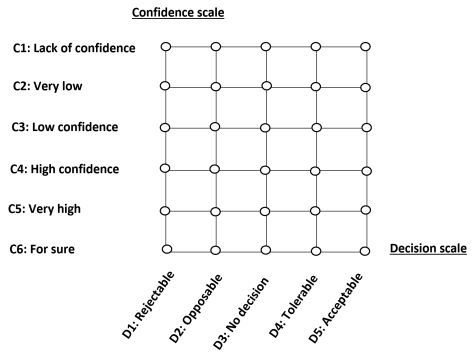


Figure A.1: Evaluation Matrix

By analogy, this is comparable to the classical conference paper reviewing where reviewers have to provide a decision (strongly reject to strongly accept) and a level of confidence associated with it (low to expert).

Examples

In the first example (figure A.2), the assessor accept the statement in goal (G1) with a very high level of confidence. This decision and its corresponding degree of confidence is based on the simulation report given as evidence (or solution in GSN nomination).

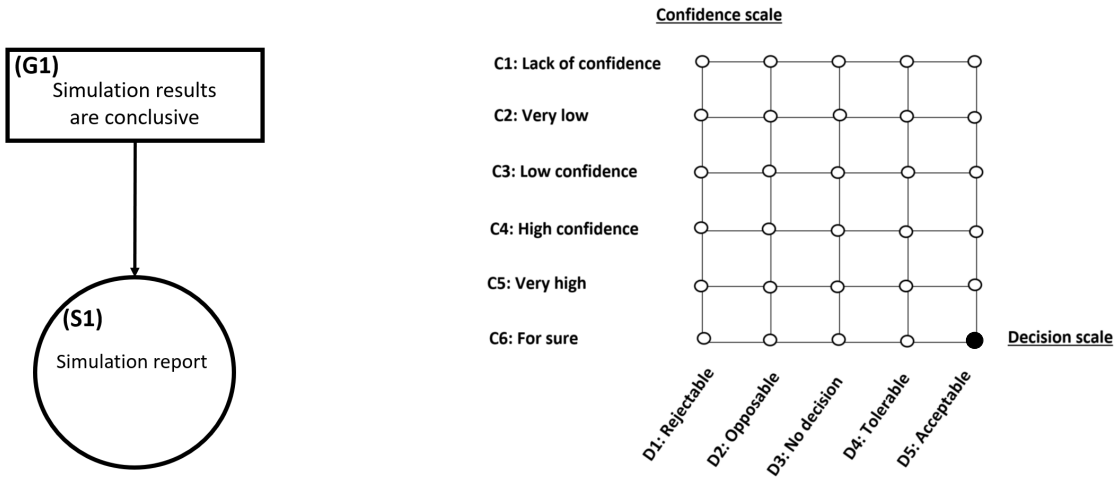


Figure A.2: Example 1

In the second example (figure A.3) the opinion about (G1) is conditioned by initial opinions on sub-goals (G2) and (G3). In this case, the assessor should give his or her decision and its corresponding level of confidence of (G1) according to the decision and confidence provided in the goals (G2) and (G3).

We can notice that the assessor is opposable to this statement with high level of confidence. This opinion could be explained by the fact that the sub-claim (G3) cannot support (G1) alone. To give a favorite opinion more information about the other sub-claim (G2) is needed.

The argument structure to assess

Figure A.4 represent an extract of the assurance case form the paper [Damour 2021], which aim to demonstrate that a Machine Learning (ML) based system meets its

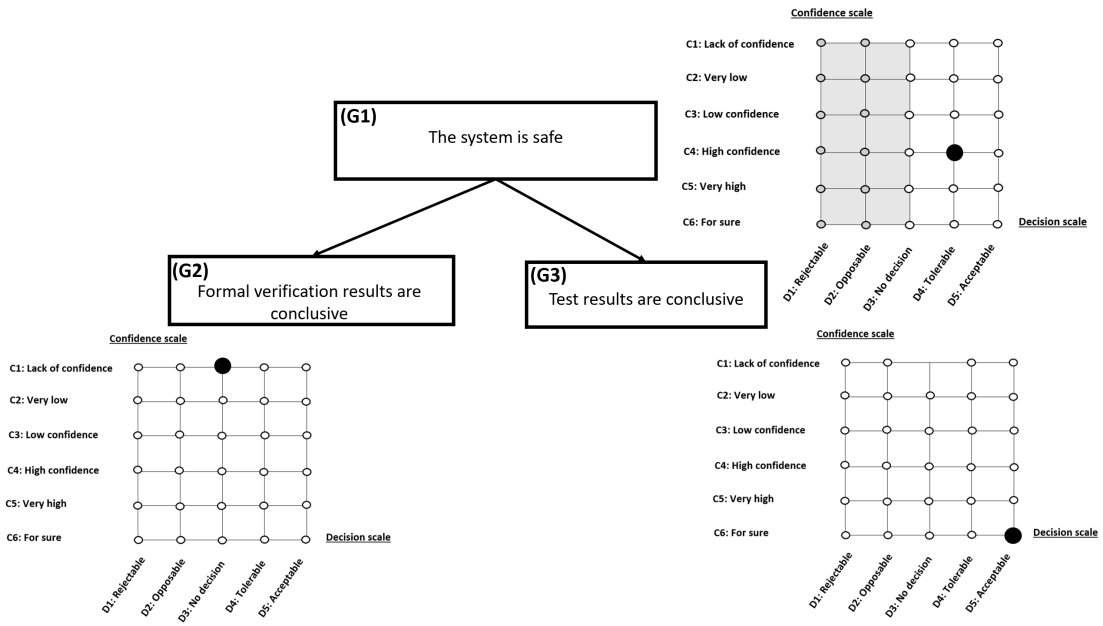


Figure A.3: Example 2

system and safety requirements. This extract shows through different pieces of evidence that the statement (G1) : "*The real world situations where the machine learning model (MLM) is not robust are identified and mitigated*" is achieved. This statement is supported by two pieces of evidence (G2) and (G3). Each of them is respectively supported by three (S1, S2 and S3) and one (S4) pieces of evidence.

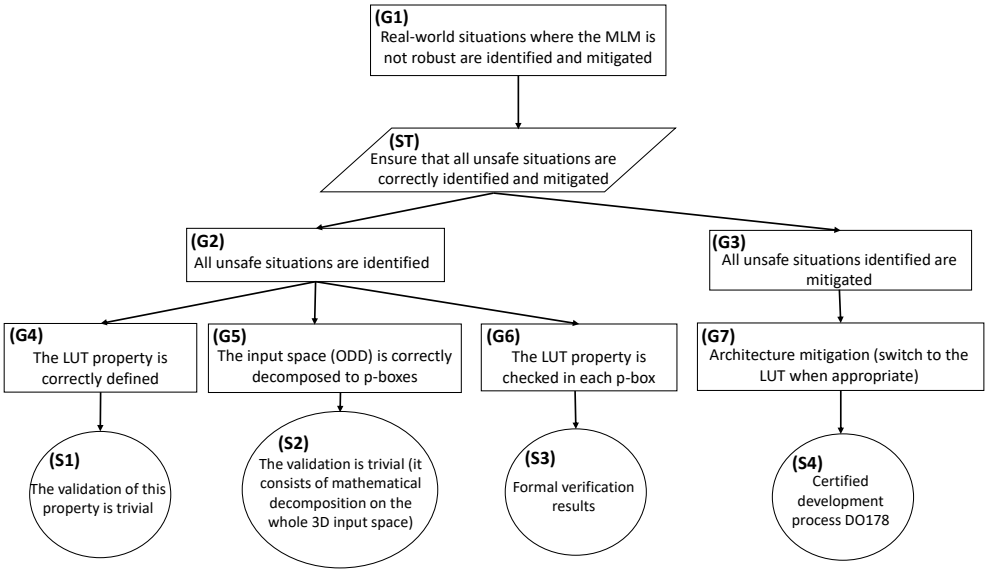
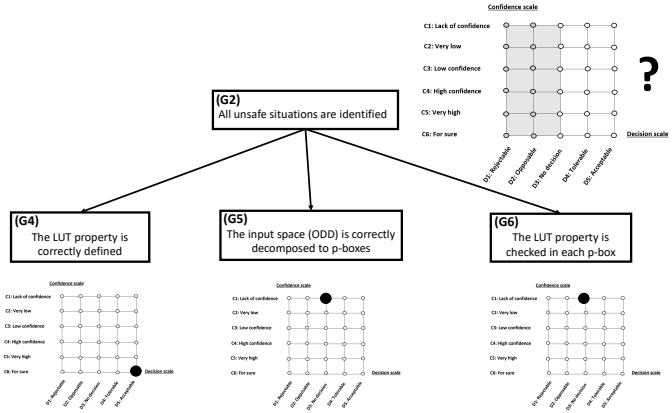
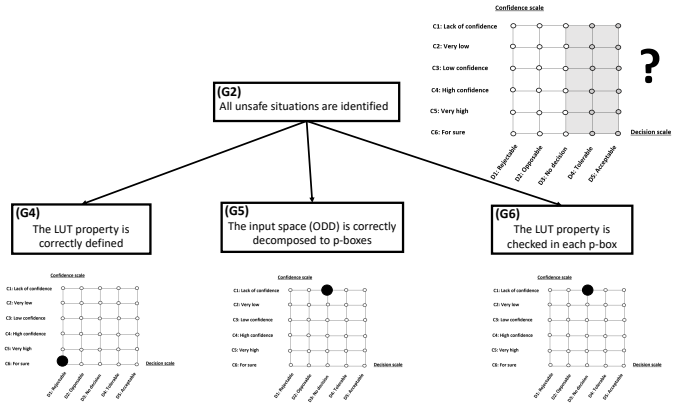


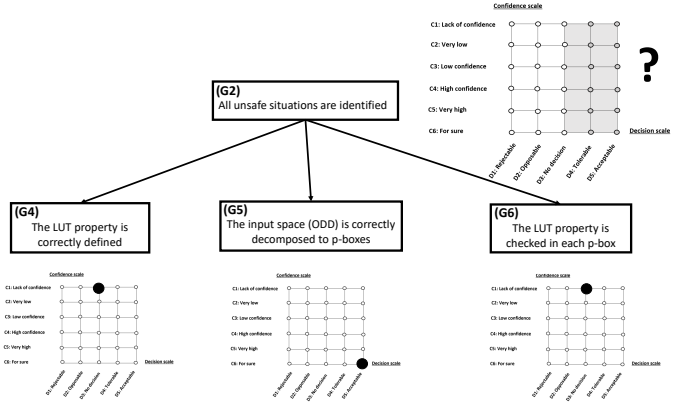
Figure A.4: Assurance Case - ML subsystem robustness [Damour 2021]

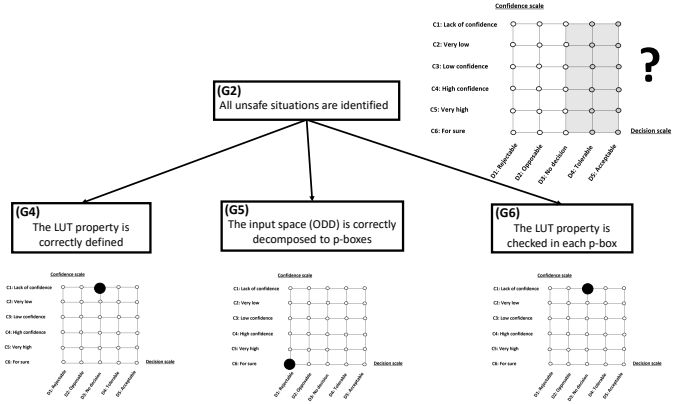
Part 1 : Elicitation of the opinions on G2

Please read the arguments bellow and give your opinion about the claim (G1) by taking into consideration the initial opinions for sub-goals (G4), (G5) and (G6). To do so, please select the couples (decision, confidence) in the goal (G2) evaluation matrix of each case. It is forbidden to choose couples in the grey area (couples situated in the median are not included).


Confidence scale
C1: Lack of confidence
C2: Very low
C3: Low confidence
C4: High confidence
C5: Very high
C6: For sure
?
Decision scale
D1: Rejectable
D2: Opposable
D3: No decision
D4: Tolerable
D5: Acceptable
(G2)
All unsafe situations are identified
(G4)
The LUT property is correctly defined
(G5)
The input space (ODD) is correctly decomposed to p-boxes
(G6)
The LUT property is checked in each p-box


Confidence scale
C1: Lack of confidence
C2: Very low
C3: Low confidence
C4: High confidence
C5: Very high
C6: For sure
?
Decision scale
D1: Rejectable
D2: Opposable
D3: No decision
D4: Tolerable
D5: Acceptable
(G2)
All unsafe situations are identified
(G4)
The LUT property is correctly defined
(G5)
The input space (ODD) is correctly decomposed to p-boxes
(G6)
The LUT property is checked in each p-box


Confidence scale
C1: Lack of confidence
C2: Very low
C3: Low confidence
C4: High confidence
C5: Very high
C6: For sure
?
Decision scale
D1: Rejectable
D2: Opposable
D3: No decision
D4: Tolerable
D5: Acceptable
(G2)
All unsafe situations are identified
(G4)
The LUT property is correctly defined
(G5)
The input space (ODD) is correctly decomposed to p-boxes
(G6)
The LUT property is checked in each p-box


Confidence scale
C1: Lack of confidence
C2: Very low
C3: Low confidence
C4: High confidence
C5: Very high
C6: For sure
?
Decision scale
D1: Rejectable
D2: Opposable
D3: No decision
D4: Tolerable
D5: Acceptable
(G2)
All unsafe situations are identified
(G4)
The LUT property is correctly defined
(G5)
The input space (ODD) is correctly decomposed to p-boxes
(G6)
The LUT property is checked in each p-box

(G2)
All unsafe situations are identified
(G4)
The LUT property is correctly defined
(G5)
The input space (ODD) is correctly decomposed to p-boxes
(G6)
The LUT property is checked in each p-box
Confidence scale
C1: Lack of confidence
C2: Very low
C3: Low confidence
C4: High confidence
C5: Very high
C6: For sure
Decision scale
D1: Rejectable
D2: Opposable
D3: No decision
D4: Tolerable
D5: Acceptable
?

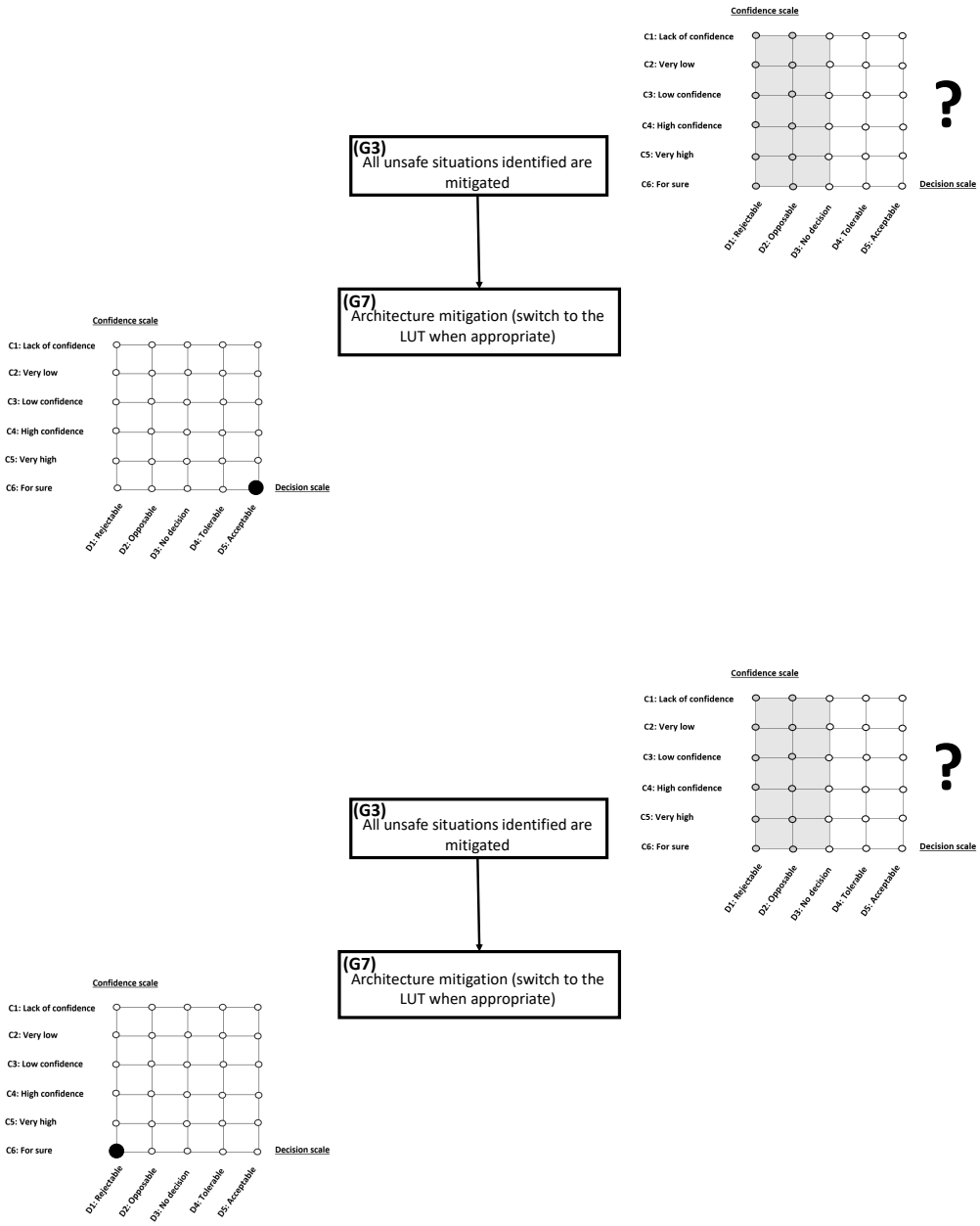
(G2)
All unsafe situations are identified
(G4)
The LUT property is correctly defined
(G5)
The input space (ODD) is correctly decomposed to p-boxes
(G6)
The LUT property is checked in each p-box
Confidence scale
C1: Lack of confidence
C2: Very low
C3: Low confidence
C4: High confidence
C5: Very high
C6: For sure
Decision scale
D1: Rejectable
D2: Opposable
D3: No decision
D4: Tolerable
D5: Acceptable
?

(G2)
All unsafe situations are identified
(G4)
The LUT property is correctly defined
(G5)
The input space (ODD) is correctly decomposed to p-boxes
(G6)
The LUT property is checked in each p-box
Confidence scale
C1: Lack of confidence
C2: Very low
C3: Low confidence
C4: High confidence
C5: Very high
C6: For sure
Decision scale
D1: Rejectable
D2: Opposable
D3: No decision
D4: Tolerable
D5: Acceptable
?

(G2)
All unsafe situations are identified
(G4)
The LUT property is correctly defined
(G5)
The input space (ODD) is correctly decomposed to p-boxes
(G6)
The LUT property is checked in each p-box
Confidence scale
C1: Lack of confidence
C2: Very low
C3: Low confidence
C4: High confidence
C5: Very high
C6: For sure
Decision scale
D1: Rejectable
D2: Opposable
D3: No decision
D4: Tolerable
D5: Acceptable
?

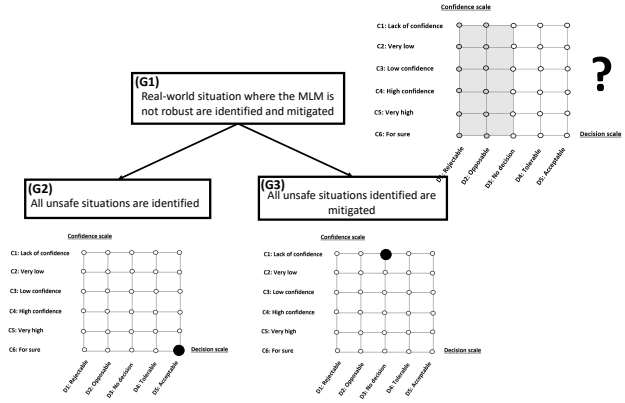
Part 2 : Elicitation of opinions on G3

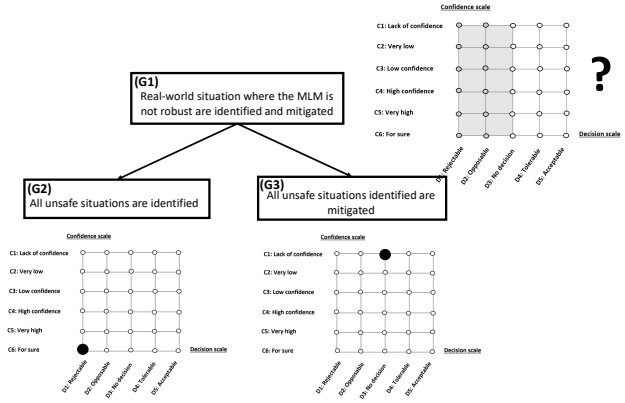
The confidence in (G3) can be deduced from the sub-goal (G7). Please read the arguments bellow and give your opinion about the goal (G3) by taking into consideration the initial opinions on the sub-goal (G7). To do so, please select the couples (decision, confidence) in the goal (G3) evaluation matrix in each case. It is forbidden to choose couples in the grey area (couples situated in the median are not included).

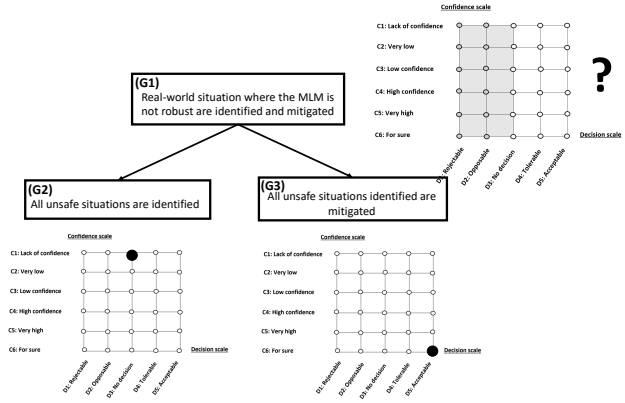


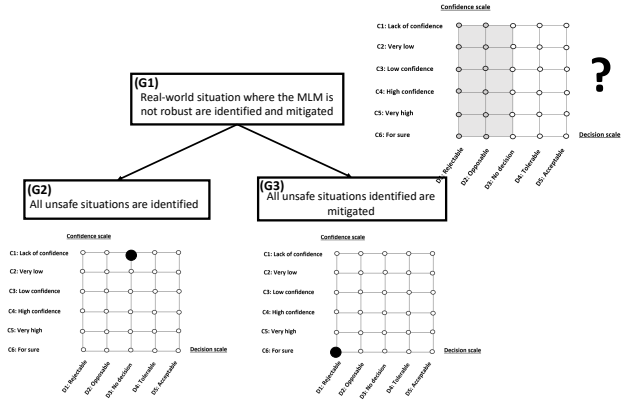
Part 3 : Elicitation of opinions on G1

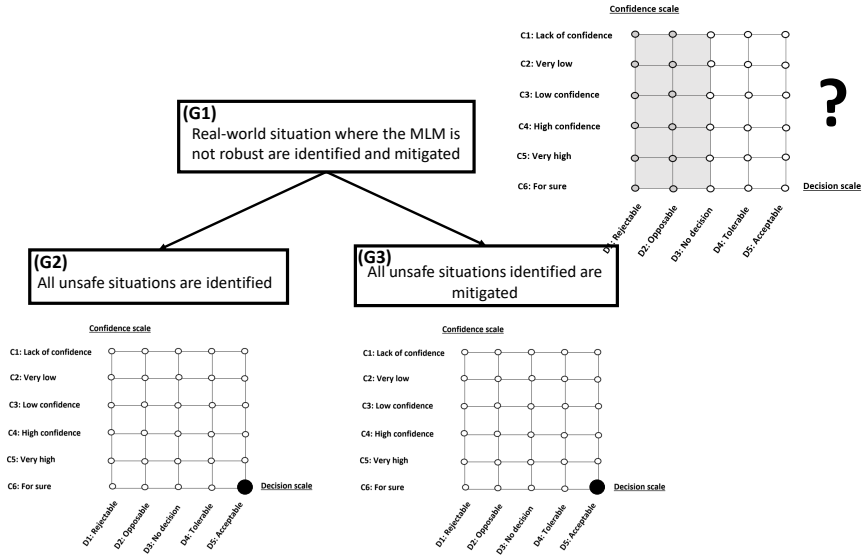
The confidence in the top goal (G1) can be deduced from these premises. Please read the arguments bellow and give your opinion about the goal (G1) by taking into consideration the initial opinions for sub-goals (G2) and (G3). To do so, please select the couples (decision, confidence) in the goal (G1) evaluation matrix in each case. It is forbidden to choose couples in the grey area (couples situated in the median are not included).

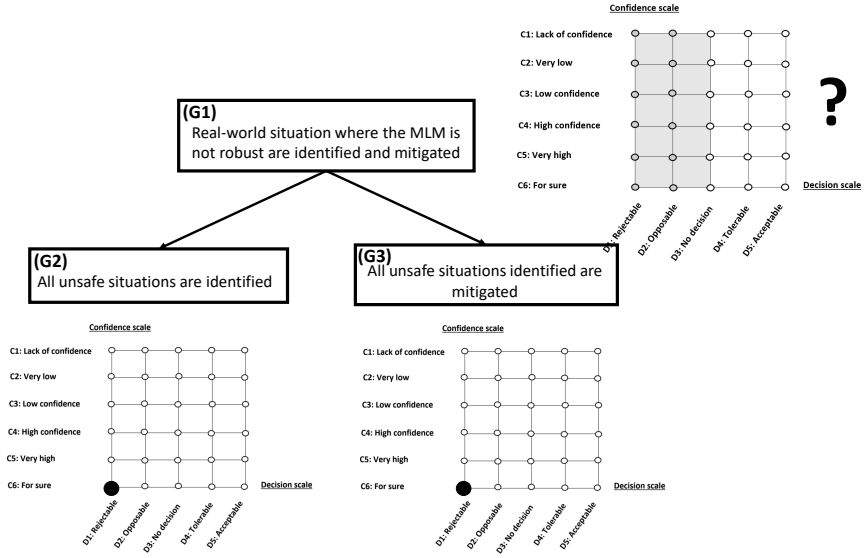

(G1)
Real-world situation where the MLM is not robust are identified and mitigated
(G2)
All unsafe situations are identified
(G3)
All unsafe situations identified are mitigated
Confidence scale
C1: Lack of confidence
C2: Very low
C3: Low confidence
C4: High confidence
C5: Very high
C6: For sure
Decision scale
D1: Rejectable
D2: Opposable
D3: No decision
D4: Tolerable
D5: Acceptable
?


(G1)
Real-world situation where the MLM is not robust are identified and mitigated
(G2)
All unsafe situations are identified
(G3)
All unsafe situations identified are mitigated
Confidence scale
C1: Lack of confidence
C2: Very low
C3: Low confidence
C4: High confidence
C5: Very high
C6: For sure
Decision scale
D1: Rejectable
D2: Opposable
D3: No decision
D4: Tolerable
D5: Acceptable
?


(G1)
Real-world situation where the MLM is not robust are identified and mitigated
(G2)
All unsafe situations are identified
(G3)
All unsafe situations identified are mitigated
Confidence scale
C1: Lack of confidence
C2: Very low
C3: Low confidence
C4: High confidence
C5: Very high
C6: For sure
Decision scale
D1: Rejectable
D2: Opposable
D3: No decision
D4: Tolerable
D5: Acceptable
?

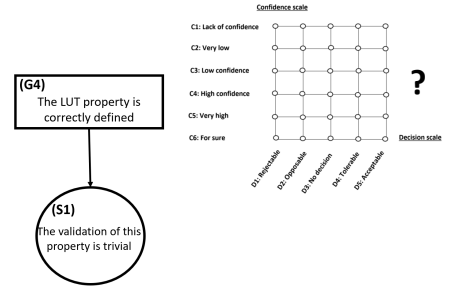

(G1)
Real-world situation where the MLM is not robust are identified and mitigated
(G2)
All unsafe situations are identified
(G3)
All unsafe situations identified are mitigated
Confidence scale
C1: Lack of confidence
C2: Very low
C3: Low confidence
C4: High confidence
C5: Very high
C6: For sure
Decision scale
D1: Rejectable
D2: Opposable
D3: No decision
D4: Tolerable
D5: Acceptable
?

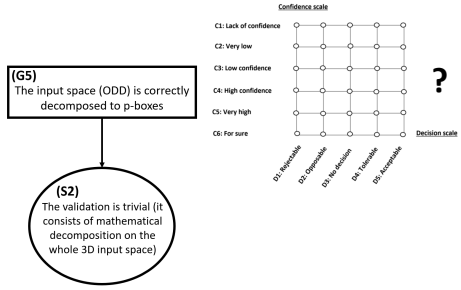

(G1)
Real-world situation where the MLM is not robust are identified and mitigated
(G2)
All unsafe situations are identified
(G3)
All unsafe situations identified are mitigated
Confidence scale
C1: Lack of confidence
C2: Very low
C3: Low confidence
C4: High confidence
C5: Very high
C6: For sure
Decision scale
D1: Rejectable
D2: Opposable
D3: No decision
D4: Tolerable
D5: Acceptable
?

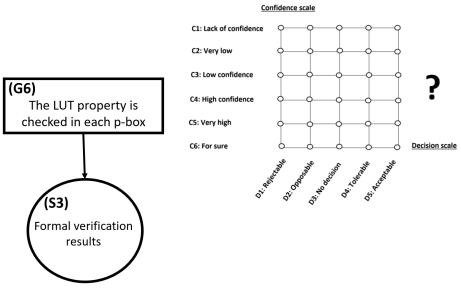

(G1)
Real-world situation where the MLM is not robust are identified and mitigated
(G2)
All unsafe situations are identified
(G3)
All unsafe situations identified are mitigated
Confidence scale
C1: Lack of confidence
C2: Very low
C3: Low confidence
C4: High confidence
C5: Very high
C6: For sure
Decision scale
D1: Rejectable
D2: Opposable
D3: No decision
D4: Tolerable
D5: Acceptable
?

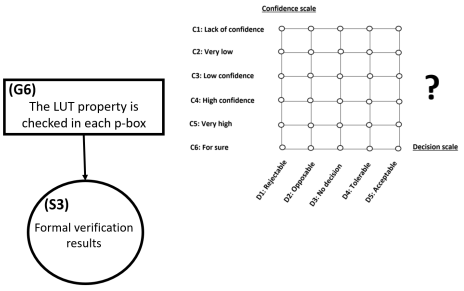
Part 4 : Elicitation of confidence and decision about premises.

In this final section, you have to give your opinion about the premises directly from the evidence provided in each argument. To do so, please read the argument below and choose the appropriate couple (decision, confidence) in each case.









Bibliography

[Ayoub 2013] A Ayoub, J Chang, O Sokolsky and I Lee. *Assessing the overall sufficiency of safety arguments.* In 21st Safety-critical Systems Symposium (SSS'13), Bristol, United Kingdom, 2013. (Cited in pages v, 19, 20, 27, 46, 47, and 82.)

[Banerjee 2014] Mohua Banerjee and Didier Dubois. *A simple logic for reasoning about incomplete knowledge.* International Journal of Approximate Reasoning, vol. 55, no. 2, pages 639–653, 2014. (Cited in page 17.)

[Bishop 2000] Peter Bishop and Robin Bloomfield. *A methodology for safety case development.* In Safety and Reliability, volume 20, pages 34–42. Taylor & Francis, 2000. (Cited in page 6.)

[Bloomfield 2010] Robin Bloomfield and Peter Bishop. *Safety and assurance cases: Past, present and possible future–an Adelard perspective.* In Making systems safer, pages 51–67. Springer, 2010. (Cited in page 10.)

[Bloomfield 2014] Robin Bloomfield and Kateryna Netkachova. *Building blocks for assurance cases.* In 2014 IEEE International Symposium on Software Reliability Engineering Workshops, pages 186–191. IEEE, 2014. (Cited in page 9.)

[Chatalic 1987a] P. Chatalic, D. Dubois and H. Prade. *An approach to approximate reasoning based on Dempster rule of combination.* Inter. J. of Expert Systems Research & Applications, vol. 1, pages 67–85, 1987. (Cited in page 17.)

[Chatalic 1987b] Philippe Chatalic, Didier Dubois and Henri Prade. *An approach to approximate reasoning based on Dempster rule of combination.* Inter. J. of Expert Systems Research & Applications, vol. 1, pages 67–85, 1987. (Cited in page 18.)

[Cholvy 2015] Laurence Cholvy. *Towards a logical belief function theory.* Journal of Applied Logic, vol. 13, no. 4, pages 441–457, 2015. (Cited in page 17.)

[Ciucci 2019] Davide Ciucci and Didier Dubois. *A capacity-based framework encompassing Belnap–Dunn logic for reasoning about multisource information.* International Journal of Approximate Reasoning, vol. 106, pages 107–127, 2019. (Cited in page 85.)

[Cobb 2006] Barry R Cobb and Prakash P Shenoy. *On the plausibility transformation method for translating belief function models to probability models.* International journal of approximate reasoning, vol. 41, no. 3, pages 314–330, 2006. (Cited in pages 17 and 65.)

[Cozman 2019] Fabio Cozman. *The joy of probabilistic answer set programming.* In International Symposium on Imprecise Probabilities: Theories and Applications, pages 91–101. PMLR, 2019. (Cited in page 17.)

[Cyra 2011] Lukasz Cyra and Janusz Górski. *Support for argument structures review and assessment.* Reliability Engineering & System Safety, vol. 96, no. 1, pages 26–37, 2011. (Cited in pages vii, 19, 20, 21, 23, 27, 29, 46, 47, 55, 56, 57, and 82.)

[Damour 2021] Mathieu Damour, Florence De Grancey, Christophe Gabreau, Adrien Gauffriau, Jean-Brice Ginestet, Alexandre Hervieu, Thomas Huraux, Claire Pagetti, Ludovic Ponsolle and Arthur Clavière. *Towards Certification of a Reduced Footprint ACAS-Xu System: A Hybrid ML-Based Solution.* In International Conference on Computer Safety, Reliability, and Security, pages 34–48. Springer, 2021. (Cited in pages vi, 9, 32, 75, 76, 77, 106, and 107.)

[De Persis 2019] Cristina De Persis, José Luis Bosque, Irene Huertas and Simon Paul Wilson. *Quantitative system risk assessment from incomplete data with belief networks and pairwise comparison elicitation.* arXiv preprint arXiv:1904.03012, 2019. (Cited in page 23.)

[de Saint-Cyr 1994] Florence Dupin de Saint-Cyr, Jérôme Lang and Thomas Schiex. *Penalty logic and its link with Dempster-Shafer theory.* In Uncertainty Proceedings 1994, pages 204–211. Elsevier, 1994. (Cited in page 17.)

[Denney 2011] Ewen Denney, Ganesh Pai and Ibrahim Habli. *Towards measurement of confidence in safety cases.* In 2011 International Symposium on Empirical Software Engineering and Measurement, pages 380–383. IEEE, 2011. (Cited in page 19.)

[Denney 2018] Ewen Denney and Ganesh Pai. *Tool support for assurance case development.* Automated Software Engineering, vol. 25, no. 3, pages 435–499, 2018. (Cited in page 9.)

[Denoeux 2008] Thierry Denoeux. *Conjunctive and disjunctive combination of belief functions induced by nondistinct bodies of evidence.* Artif. Intell., vol. 172, no. 2-3, pages 234–264, 2008. (Cited in page 16.)

[Destercke 2011] Sébastien Destercke and Didier Dubois. *Idempotent conjunctive combination of belief functions: Extending the minimum rule of possibility theory.* Inf. Sci., vol. 181, no. 18, pages 3925–3945, 2011. (Cited in page 16.)

[Duan 2014] Lian Duan, Sanjai Rayadurgam, Mats PE Heimdahl, Anaheed Ayoub, Oleg Sokolsky and Insup Lee. *Reasoning about confidence and uncertainty in assurance cases: A survey.* Software Engineering in Health Care, pages 64–80, 2014. (Cited in pages 9 and 10.)

[Dubois 1988] Didler Dubois and Henri Prade. *Representation and combination of uncertainty with belief functions and possibility measures.* Computational intelligence, vol. 4, no. 3, pages 244–264, 1988. (Cited in page 12.)

[Dubois 2019] Didier Dubois, Francis Faux, Henri Prade and Agnès Rico. *A possibilistic counterpart to Shafer evidence theory.* In 2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pages 1–6. IEEE, 2019. (Cited in page 84.)

[Dubois 2022a] Didier Dubois, Francis Faux, Henri Prade and Agnes Rico. *Qualitative capacities: Basic notions and potential applications.* Int. J. of Approximate Reasoning, vol. in press, 2022. (Cited in page 84.)

[Dubois 2022b] Didier Dubois, Francis Faux, Henri Prade and Agnès Rico. *Qualitative capacities: Basic notions and potential applications.* International Journal of Approximate Reasoning, 2022. (Cited in page 92.)

[Gallina 2016] Barbara Gallina, Elena Gómez-Martínez and Clara Benac Earle. *Deriving safety case fragments for assessing mbasafe's compliance with en 50128.* In International Conference on Software Process Improvement and Capability Determination, pages 3–16. Springer, 2016. (Cited in page 9.)

[Gauerhof 2020] Lydia Gauerhof, Richard Hawkins, Chiara Picardi, Colin Paterson, Yuki Hagiwara and Ibrahim Habli. *Assuring the safety of machine learning for pedestrian detection at crossings.* In International Conference on Computer Safety, Reliability, and Security, pages 197–212. Springer, 2020. (Cited in page 9.)

[Górski 2005a] Janusz Górski. *Trust case—a case for trustworthiness of IT infrastructures.* In Cyberspace Security and Defense: Research Issues, pages 125–141. Springer, 2005. (Cited in page 6.)

[Górski 2005b] Janusz Górski, Aleksander Jarzębowicz, Rafal Leszczyna, Jakub Miler and Marcin Olszewski. *Trust case: Justifying trust in an IT solution.* Reliability Engineering & System Safety, vol. 89, no. 1, pages 33–47, 2005. (Cited in page 6.)

[Graydon 2017] Patrick J Graydon and C Michael Holloway. *An investigation of proposed techniques for quantifying confidence in assurance arguments.* Safety science, vol. 92, pages 53–65, 2017. (Cited in pages 10 and 27.)

[Guiochet 2014] Jérémie Guiochet, Quynh Anh Do Hoang and Mohamed Kaaniche. *A model for safety case confidence assessment.* In International Conference on Computer Safety, Reliability, and Security (Safecomp), pages 313–327. Springer, 2014. (Cited in pages 19 and 82.)

[Habli 2010] Ibrahim Habli, Ireri Ibarra, Roger Rivett and Tim Kelly. *Model-based assurance for justifying automotive functional safety.* In Proc. 2010 SAE World Congress, volume 1, 2010. (Cited in page 9.)

[Harmanec 1999] David Harmanec. *Measures of uncertainty and information.* Imprecise Probability Project, 1999. (Cited in page 11.)

[Hawkins 2021] Richard Hawkins, Colin Paterson, Chiara Picardi, Yan Jia, Radu Calinescu and Ibrahim Habli. *Guidance on the assurance of machine learning in autonomous systems (AMLAS).* arXiv preprint arXiv:2102.01564, 2021. (Cited in page 9.)

[Hobbs 2012] Chris Hobbs and Martin Lloyd. *The application of bayesian belief networks to assurance case preparation.* In Achieving Systems Safety, pages 159–176. Springer, 2012. (Cited in page 19.)

[Idmessaoud 2020] Yassir Idmessaoud, Didier Dubois and Jérémie Guiochet. *Belief functions for safety arguments confidence estimation: A comparative study.* In International Conference on Scalable Uncertainty Management, pages 141–155. Springer, 2020. (Cited in pages 28 and 105.)

[Idmessaoud 2021a] Yassir Idmessaoud, Didier Dubois and Jérémie Guiochet. *Quantifying confidence of safety cases with belief functions.* In International Conference on Belief Functions, pages 269–278. Springer, 2021. (Cited in pages 28 and 105.)

[Idmessaoud 2021b] Yassir Idmessaoud, Jérémie Guiochet and Didier Dubois. *Calculation of aggregation formulas for GSN argument types using belief functions.* 2021. (Cited in page 105.)

[Idmessaoud 2022] Yassir Idmessaoud, Didier Dubois and Jérémie Guiochet. *A qualitative counterpart of belief functions with application to uncertainty propagation in safety cases.* In International Conference on Belief Functions, pages 231–241. Springer, 2022. (Cited in page 105.)

[Inagaki 1991] Toshiyuki Inagaki. *Interdependence between safety-control policy and multiple-sensor schemes via Dempster-Shafer theory.* IEEE Transactions on Reliability, vol. 40, no. 2, pages 182–188, 1991. (Cited in page 16.)

[ISO 26262 2011] ISO 26262. *Software Considerations in Airborne Systems and Equipment Certification*, 2011. International Organization for Standardization (ISO). (Cited in pages 6 and 10.)

[Jiroušek 2014] Radim Jiroušek and Prakash P Shenoy. *Compositional models in valuation-based systems.* International journal of approximate reasoning, vol. 55, no. 1, pages 277–293, 2014. (Cited in page 18.)

[Jøsang 2016] Audun Jøsang. Subjective logic. Springer, 2016. (Cited in page 59.)

[Katz 2017] Guy Katz, Clark Barrett, David L Dill, Kyle Julian and Mykel J Kochenderfer. *Reluplex: An efficient SMT solver for verifying deep neural networks.* In International conference on computer aided verification, pages 97–117. Springer, 2017. (Cited in pages vi and 76.)

[Kelly 1997] Tim P Kelly and John A McDermid. *Safety case construction and reuse using patterns.* In Safe Comp 97, pages 55–69. Springer, 1997. (Cited in pages v and 8.)

[Kelly 2004] Tim Kelly and Rob Weaver. *The goal structuring notation–a safety argument notation.* In Proceedings of the dependable systems and networks 2004 workshop on assurance cases, page 6. Citeseer, 2004. (Cited in pages v, 6, 7, and 9.)

[Leppäaho 2021] Oskari Leppäaho, Mark Nicholson, Frédéric Lafon and Mohammed Ramdani. *Assuring Shielded Cables as EMI Mitigation in Automotive ADAS.* In 31st European Safety and Reliability Conference, pages 1859–1866. Research Publishing Services, 2021. (Cited in page 9.)

[Matsuno 2010] Yutaka Matsuno, Hiroki Takamura and Yutaka Ishikawa. *A dependability case editor with pattern library.* In 2010 IEEE 12th International Symposium on High Assurance Systems Engineering, pages 170–171. IEEE, 2010. (Cited in pages 6 and 9.)

[MoD 00-56 2007] MoD 00-56. *Defence Standard 00-56 Issue 4 (Part 1): Safety Management Requirements for Defence Systems.* UK Ministry of Defence, June 1, 2007. (Cited in page 7.)

[Nair 2015] Sunil Nair, Neil Walkinshaw, Tim Kelly and Jose Luis de la Vara. *An evidential reasoning approach for assessing confidence in safety evidence.* In 2015 IEEE 26th International Symposium on Software Reliability Engineering (ISSRE), pages 541–552. IEEE, 2015. (Cited in page 23.)

[Nešić 2021] Damir Nešić, Mattias Nyberg and Barbara Gallina. *A probabilistic model of belief in safety cases.* Safety science, vol. 138, page 105187, 2021. (Cited in page 19.)

[Palin 2010] Robert Palin and Ibrahim Habli. *Assurance of automotive safety–a safety case approach.* In International Conference on Computer Safety, Reliability, and Security, pages 82–96. Springer, 2010. (Cited in page 9.)

[Picardi 2020] Chiara Picardi, Colin Paterson, Richard David Hawkins, Radu Calinescu and Ibrahim Habli. *Assurance argument patterns and processes for machine learning in safety-related systems.* In Proceedings of the Workshop on Artificial Intelligence Safety (SafeAI 2020), pages 23–30. CEUR Workshop Proceedings, 2020. (Cited in page 9.)

[Porter 2022] Zoe Porter, Ibrahim Habli and John McDermid. *A Principle-based Ethical Assurance Argument for AI and Autonomous Systems.* arXiv preprint arXiv:2203.15370, 2022. (Cited in page 9.)

[Rushby 2015] John Rushby. *The interpretation and evaluation of assurance cases.* Comp. Science Laboratory, SRI International, Tech. Rep. SRI-CSL-15-01, 2015. (Cited in page 6.)

[Selviandro 2020] Nungki Selviandro, Richard Hawkins and Ibrahim Habli. *A visual notation for the representation of assurance cases using sacm.* In International Symposium on Model-Based Safety and Assessment, pages 3–18. Springer, 2020. (Cited in page 9.)

[Sentz 2002] Kari Sentz, Scott Ferson*et al. Combination of evidence in Dempster-Shafer theory.* Technical Report 0835, Sandia National Laboratories, Albuquerque, NM., USA, 2002. (Cited in pages 14 and 16.)

[Shafer 1976] Glenn Shafer. A mathematical theory of evidence. Princeton university press, 1976. (Cited in pages 12, 14, 15, and 29.)

[Shenoy 1989] Prakash P Shenoy. *A valuation-based language for expert systems.* International Journal of Approximate Reasoning, vol. 3, no. 5, pages 383–411, 1989. (Cited in page 18.)

[Shenoy 1990] Prakash P. Shenoy and Glenn Shafer. *Axioms for Probability and Belief-Function Propagation.* In Ross D. Shachter, Tod S. Levitt, Laveen N. Kanal and John F. Lemmer, editors, Uncertainty in Artificial Intelligence, volume 9 of *Machine Intelligence and Pattern Recognition*, pages 169–198. North-Holland, 1990. (Cited in page 18.)

[Smets 2005] P. Smets. *Decision making in the TBM: the necessity of the pignistic transformation.* International Journal of Approximate Reasoning, vol. 38, pages 133–147, 2005. (Cited in pages 16, 17, and 56.)

[Smets 2007] Philippe Smets. *Analyzing the combination of conflicting belief functions.* Inf. Fusion, vol. 8, no. 4, pages 387–412, 2007. (Cited in page 14.)

[Toulmin 1958] S. E. Toulmin. The uses of argument. Cambridge Univ. Press, Cambridge, 1958. (Cited in page 7.)

[Walley 1991] Peter Walley. *Statistical reasoning with imprecise probabilities.* 1991. (Cited in page 14.)

[Wang 2016] Rui Wang, Jérémie Guiochet, Gilles Motet and Walter Schön. *DS theory for argument confidence assessment.* In International Conference on Belief Functions, pages 190–200. Springer, 2016. (Cited in page 17.)

[Wang 2018a] Rui Wang. *Confidence in safety argument-An assessment framework based on belief function theory.* PhD thesis, Institut National des Sciences Appliquées de Toulouse, France, 2018. (Cited in pages 21, 23, 28, 46, 47, 55, 82, and 101.)

[Wang 2018b] Rui Wang, Jérémie Guiochet, Gilles Motet and Walter Schön. *Modelling Confidence in Railway Safety Case.* Safety Science, vol. 110 part B, pages 286–299, December 2018. (Cited in pages v, 21, 23, 46, 56, and 64.)

[Wang 2019] Rui Wang, Jérémie Guiochet, Gilles Motet and Walter Schön. *Safety Case Confidence Propagation Based on Dempster-Shafer theory.* International Journal of Approximate Reasoning, vol. 107, pages 46–64, April 2019. (Cited in pages vii, 17, 19, 22, 27, 28, 29, 46, 47, 57, and 100.)

[Wei 2019] Ran Wei, Tim P Kelly, Xiaotian Dai, Shuai Zhao and Richard Hawkins. *Model based system assurance using the structured assurance case metamodel.* Journal of Systems and Software, vol. 154, pages 211–233, 2019. (Cited in page 9.)

[Weinstock 2004] Charles B Weinstock, John B Goodenough and John J Hudak. *Dependability cases.* Technical Report, CARNEGIE-MELLON UNIV PITTSBURGH PA SOFTWARE ENGINEERING INST, 2004. (Cited in page 6.)

[Weinstock 2013] Charles B Weinstock, John B Goodenough and Ari Z Klein. *Measuring assurance case confidence using Baconian probabilities.* In 2013 1st international workshop on assurance cases for software-intensive systems (ASSURE), pages 7–11. IEEE, 2013. (Cited in page 19.)

[Yager 1987] Ronald R Yager. *On the Dempster-Shafer framework and new combination rules.* Information sciences, vol. 41, no. 2, pages 93–137, 1987. (Cited in pages 15 and 16.)

[Yaghlane 2008] Amel Ben Yaghlane, Thierry Denœux and Khaled Mellouli. *Elicitation of expert opinions for constructing belief functions.* In Uncertainty and intelligent information systems, pages 75–89. World Scientific, 2008. (Cited in page 23.)

[Yuan 2017] Chunchun Yuan, Ji Wu, Chao Liu and Haiyan Yang. *A subjective logic-based approach for assessing confidence in assurance case.* International Journal of Performability Engineering, vol. 13, no. 6, pages 807–822, 2017. (Cited in page 19.)

[Zadeh 1999] Lotfi A Zadeh. *Fuzzy sets as a basis for a theory of possibility.* Fuzzy sets and systems, vol. 100, pages 9–34, 1999. (Cited in pages 12 and 84.)

Abstract:

Structured assurance cases are use to justify high-level properties (e.g. safety, security, etc.) of critical systems. Goal structuring notation (GSN) is a graphical notation used to model these cases. However, assurance cases do 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. Moreover, the approaches used to provide inputs for these propagation models rely on expert judgements to assess the arguments. These assessments are collected in natural languages and transformed to numerical values. This transformation is somewhat arbitrary. In this thesis, we first improve previous proposals including a logical formalism added to GSN, and an elicitation procedure for obtaining uncertainty information from expert judgements. Next, we propose a second confidence assessment approach which is based on a purely qualitative propagation model fed by a qualitative elicitation method too. These approaches are validated using artificial examples and a case study from aerospace.

Keywords: Assurance cases, Safety cases, Dempster-Shefer Theory (DST), Confidence propagation, Expert elicitation, Goal Structuring Notation (GSN), Qualitative Capacity Theory (QCT).