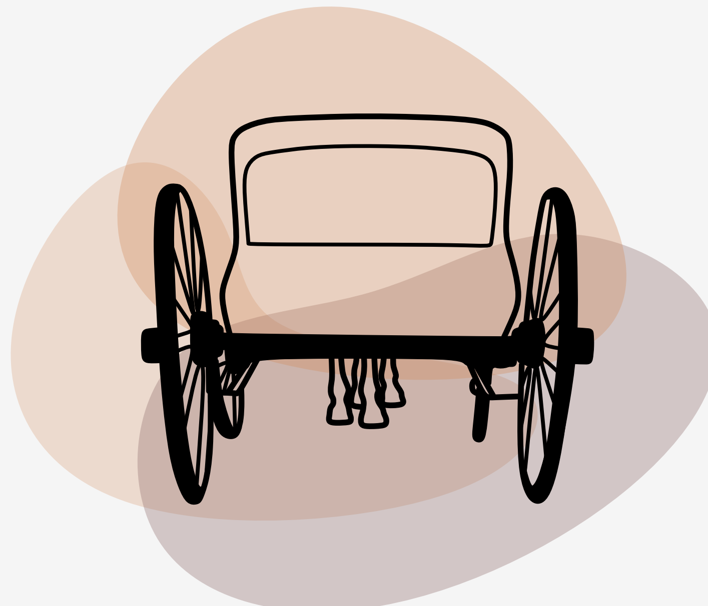
Combining Uncertain Evidence

Logic and Complexity



Daira Pinto Prieto

Combining Uncertain Evidence Logic and Complexity

Daira Pinto Prieto

Institute for Logic, Language and Computation

Combining Uncertain Evidence

Logic and Complexity

Daira Pinto Prieto

Combining Uncertain Evidence

Logic and Complexity

ILLC Dissertation Series DS-2024-11

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

For further information about ILLC-publications, please contact

Institute for Logic, Language and Computation
Universiteit van Amsterdam
Science Park 107
1098 XG Amsterdam
phone: +31-20-525 6051
e-mail: illc@uva.nl
homepage: http://www.illc.uva.nl/

Cover design by Arthur Boixel and Daira Pinto Prieto

Printed and bound by Ipskamp Printing

ISBN: 978–94–6473–618–2

Combining Uncertain Evidence
Logic and Complexity

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. P.P.C.C. Verbeek
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op maandag 25 november 2024, te 14.00 uur

door Daira Pinto Prieto
geboren te León

Promotiecommissie

Promotor:	prof. dr. U. Endriss	Universiteit van Amsterdam
Copromotores:	dr. R. de Haan	Universiteit van Amsterdam
	dr. A. Özgün	Universiteit van Amsterdam
Overige leden:	prof. dr. S.J.L. Smets	Universiteit van Amsterdam
	prof. dr. Y. Venema	Universiteit van Amsterdam
	dr. B.D. ten Cate	Universiteit van Amsterdam
	prof. dr. M. Slavkovik	University of Bergen
	dr. S. Destercke	Université de Technologie de Compiègne

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

A mis abuelos: Ascensión, Félix, Lola y Ramiro.

Contents

Acknowledgments

This dissertation is the result of four years of work. Beyond my own efforts, its completion would not have been possible without the contributions of many different people. I would like to thank those who have given me the guidance, the energy, and the opportunity to reach this milestone.

First, regarding guidance, I would like to thank my supervisor, Ronald de Haan, for his continuous encouragement and positive attitude towards all my initiatives. I also thank Aybüke Özgün, who has co-supervised this project for the past three years, for her enthusiasm and meticulous contributions throughout the process. It is easier to thrive academically when you have the full trust and advice of such brilliant colleagues.

Next, concerning energy, I would like to thank Arthur Boixel for ensuring my well-being during this intensive period of work. Thank you for preparing meals, keeping our living space clean, reminding me to take breaks, and motivating me to get to work when I struggled to start. It is a privilege to engage in a job that requires long hours of focus without sacrificing many other aspects of life; thank you for teaming up with me to make it possible. Thank you too for bringing my design wishes to life, including the formatting and the cover of this book.

Finally, I would like to thank my promotor, Ulle Endriss, for the opportunity to become a researcher, as well as my doctoral committee members—Balder ten Cate, Sébastien Destercke, Marija Slavkovik, Sonja Smets, and Yde Venema—for agreeing to evaluate this dissertation. I am also grateful to the University of Amsterdam and the Institute for Logic, Language and Computation for their financial and administrative support.

Now, finishing something is one thing, but how you finish it—emotionally speaking—is another. I feel immensely lucky to be surrounded by people who love me, who are genuinely happy when things go well for me, and who are always ready to lend a hand when they do not. I also want to thank them all for making my journey to and through my PhD a joyful one.

While writing these lines, I cannot help but think about the people who have taken care of me during my academic career. All the years of study leading to this point have been built on collective support; partly from society through various scholarships, and most importantly from my family. My first thank-you in this regard is to them: to my mother for always being available; to my father for always finding a solution; to my grandparents for sharing everything they had; to my brothers and sister for always having the right meme ready; to my partner for being home in every sense of the word; and to my aunts and uncles for accompanying that little girl who grew up and is here today.

My second thank-you is for those who have discovered the world alongside me—my friends. The years pass, but I carry with me the memories of our walks around the playground, the endless notes we exchanged in class, the hours in the library and *delegación*, the parties, the trips, and the shared coffees. I will always hold León, Oviedo, and Madrid in my heart, and now also Amsterdam and the friends it has brought me. To them, thank you for becoming my favorite reason to come to the office. PhD life is too unpredictable to regularly get together, but after disappearing for travels or peaks of work, it is truly comforting to reconnect over long conversations, coffees, and dinners.

In short, thank you to everyone who has laughed, shared endless chats, and enjoyed tasty food with me. With such good company, it is easy to find the strength to accomplish anything.

Daira Pinto Prieto
Amsterdam, September 2024

Chapter 1

Introduction

Nowadays, we are extremely used to understanding the world through data. The news explain how the planet will change according to scientific evidence; social media show us advertisements based on our likes and searches; and our smartphones provide us with an overview of our fitness health based on the number of steps and hours of sleep we had in the day. These are just three examples among hundreds that illustrate our interaction with modern life: heterogeneous data is collected, some kind of math is used to aggregate it, and a conclusion is presented to us. This thesis is about that *kind of math*.

Before delving into the specific context we study, let us take a moment to consider the big picture. First, the data (or information) we gather. As mentioned, it is heterogeneous. It may originate from sensors or human interactions, among others. It may come from a single source or a collection of sources. It may be certain or doubtful, coherent or not. Gathering raw information from the world (instead of from a controlled environment like a scientific lab) opens the door to collecting any kind of information, in any format, and with any level of reliability. This leads to a daunting realization for a mathematician: we can assume very little about the input of our problem. Second, the math used to aggregate this data and provide a conclusion must rely on a solid theoretical foundation. The reason is simple: most of the time, this math is invisible to users, sometimes even untrackable for developers (so-called *black boxes*). Moreover, as we mentioned, these processes are everywhere. So we must ensure that these mistakes occur only exceptionally (and continue researching to avoid them). Last but not least, the conclusion that the whole process presents will be perceived as objective information because *the data says it all*. Therefore, we have an additional argument to conduct theoretical research in information fusion.

In our case, we will not talk about information fusion in general. We will focus on merging pieces of evidence that can be represented as sets of elements within a finite universe of elements. A piece of evidence will be understood as an observation made or obtained from a source in the real world. For example, imagine Player 1 is playing *Guess Who?* against Player 2. The solution universe for Player 1 will be the collection of characters that are on their board, let us say $\{\text{Ami}, \text{Taylor}, \text{Gabe}, \text{Holly}, \text{Casey}, \text{Sally}\}$. Now, if Player 1 asks, "Is this person wearing a hat?" and Player 2 answers, "No", Player 1 has a piece of evidence for the set of characters that are not wearing a hat. If these characters are Taylor, Gabe, Casey and Sally, this piece of evidence would be represented as

$$\{\text{Taylor}, \text{Gabe}, \text{Casey}, \text{Sally}\}$$

Therefore, we do not consider information such as text or images, at least not without preprocessing it and representing it as sets of elements. In other words, we represent evidence as sets of possible worlds. In epistemology this kind of information is often called propositional evidence. That is why we talk about combining evidence instead of information fusion.

As mentioned, we aim to aggregate heterogeneous evidence, so we will introduce some "imperfections" in our representation. First, we will assume that our pieces of evidence have some degree of uncertainty. We will represent the degree of uncertainty of a piece of evidence by a number between 0 and 1. For instance, imagine that Player 1 asks, "Is this person's name typically feminine?", Player 2 does not recognize all the names on the board and they are allowed to give a certainty degree together with their yes/no answer. Then, if the answer is "No, but my certainty degree is only 0.6",

$$(\{\text{Taylor}, \text{Gabe}, \text{Casey}\}, 0.6)$$

would represent an uncertain piece of evidence. We will also allow having partial information. That is, if Player 2 tells Player 1 that the character is blonde with a certainty of 0.75, Player 1 ignores which is the certainty for having a character with black, brown, or red hair, only attending to the available evidence. Additionally, we will assume that the pieces of evidence can be mutually contradictory, which we will represent as sets with empty intersections. For instance, a set of pieces of evidence

$$\Big\{(\{\text{Ami}, \text{Taylor}, \text{Gabe}\}, 0.8), (\{\text{Holly}, \text{Casey}, \text{Sally}\}, 0.65)\Big\}$$

is a valid input for the problem. With this game in mind only, this much flexibility might seem unwarranted. But think about a legal case where the judge is trying to find a culprit or culprits of an infraction. It is perfectly possible to gather mutually contradictory evidence from independent sources with apparently incoherent degrees of certainty—i.e., one source might provide evidence A and the other the set complement (negation) of A with the same degree of certainty, e.g., 0.8. Furthermore, this evidence will be partial. Partial evidence in this context means that we may not have the information to decide for or against each individual defendant. In particular, we know that in a court of law, to have evidence of the innocence of one person does not mean to have evidence of the guilt of the others. This is the kind of relationship between pieces of evidence that we will have in our representation.

Apart from these permissive assumptions, we will make a restrictive one. We assume that all the pieces of evidence come from independent sources. However, we will point to related work that drops this restriction as well.

Once we have this uncertain, partial, and possibly mutually contradictory evidence, our goal is to combine it and draw a conclusion based on it. In our case, this conclusion will be computing degrees of belief. So we want Player 1 to be able to answer the question "What is my degree of belief in ...?", where the dots can be replaced by any set of the game's candidates—which we will call *proposition*. Wrapping up all this information, our informal research question is: *How to merge observations from the real world to draw conclusions that are guaranteed to be consistent?* Using more technical terms, this rephrases as

How to combine uncertain, partial and possibly mutually contradictory evidence and compute degrees of belief based on it?

We will introduce two well-established mathematical frameworks that answer this question, namely, Dempster-Shafer theory and topological models of evidence, in Chapter 2. These frameworks are rather complementary, since the former is quantitative and the latter is purely qualitative. In Chapter 3 we will propose our own solution, which aims to combine the strengths of both to provide a richer framework. In particular, we define a model to compute the degree of belief based on uncertain evidence, which counts on two parameters closely related to the topological models of evidence. By changing the values of these parameters, our belief model can produce the same results as some well-known approaches—concretely the two we take as base cases—as well as unexplored ones. As the subtitle of this thesis claims, we will address this problem from a logic and a computational complexity perspective. Chapter 4 is dedicated to the former. We will define an epistemic logic with two modal operators that compare propositions, one according to the certainty of the evidence supporting them and another according to their degree of belief. We will give some initial steps to discover the generated logic. The computational complexity of the problem is studied in Chapters 5 and 6. Chapter 5 is dedicated to understanding the computational complexity of the proposed solutions. Chapter 6 presents a concrete algorithmic solution to apply the (main) combination tool of Dempster-Shafer theory in an efficient-enough way in certain contexts.

This thesis is the result of analyzing our research question from different angles, from its mathematical foundations to epistemic logic and computational complexity. We hope it leads to interesting research lines and encourages further collaborations.

Chapter Overview

Chapter 2: Background In this chapter, we clarify the scope of this thesis and introduce the elements of Dempster-Shafer theory and topological models of evidence that will serve as cornerstones of our work. In Sections 2.1 and 2.2, readers can respectively familiarize themselves with Dempster-Shafer theory and topological models of evidence within the context of their respective fields and research questions. In Section 2.3, we highlight the similarities between these frameworks and how they complement each other. We also establish a common vocabulary for the rest of this thesis. We close the chapter with a clarification table that compares our terminology with those of Dempster-Shafer theory and topological models of evidence. Part of the content of this chapter is based on the preliminaries section of the paper (Pinto Prieto et al., 2023).

Chapter 3: Multi-Layer Belief Model for Combining Uncertain Evidence The focus of this chapter is to define a mathematical model that generalizes both Dempster-Shafer theory and topological models of evidence (under some restrictions). As a result, we enrich both frameworks by adding quantitative components to topological models of evidence and by introducing different notions of evidence into Dempster-Shafer theory. We introduce this method to combine uncertain evidence as the *multi-layer belief model* because it is defined in three separate steps or layers. In Section 3.1, we define these three layers. The first layer focuses on the sets that represent pieces of evidence; it is intimately connected to topological models of evidence; and introduces the notion of the evidential demand of the agent. That is, different agents may establish different requirements for the available evidence to be admissible to support their beliefs. The second layer aims to combine the degrees of uncertainty of the various pieces of evidence. It is deeply inspired by Dempster-Shafer theory. The last layer connects the previous two and finalizes the process of obtaining degrees of belief. To do so, the agents can set a context parameter related to how reliable the sources are. In Section 3.2, we prove that, by setting the evidential demand and the context parameter in certain ways, the multi-layer belief model manages to reproduce some combination rules of Dempster-Shafer theory and the belief operator of topological models of evidence, as well as introduce new alternatives. Section 3.3 presents a real-life scenario where the versatility of the multi-layer belief model could be insightful. This chapter reproduces and extends the content of the paper (Pinto Prieto et al., 2023).

Chapter 4: A Qualitative Logic for Evidence and Belief Comparison This chapter is a first step towards developing a modal logic for the multi-layer belief model. In particular, we propose a qualitative logic that explicitly compares the strengths of beliefs and evidential support with respect to the certainty of the available evidence. The latter is the novel component of this logic, so we dedicate Section 4.2 to motivating and justifying our definition of a modal operator that compares propositions according to the degree of certainty of the evidential support they receive. In Section 4.3, we define the syntax, the formal models, and the semantics of our logic. We also explore some of its validities and invalidities. Section 4.4 is dedicated to presenting a real-life scenario where being able to reason about evidence uncertainty and the resulting belief function could be useful. The content of this chapter is based on an unpublished manuscript co-authored with Aybüke Özgün (Pinto Prieto and Özgün, 2024).

Chapter 5: Computational Complexity of Belief based on Evidence This chapter begins a systematic analysis of the computational complexity of three problems: applying Dempster's rule of combination (one of the combination rules of Dempster-Shafer theory), obtaining the outcome of the belief operator defined in topological models of evidence, and applying the multi-layer belief model. Section 5.1 is based on the paper (Pinto Prieto and De Haan, 2022). It addresses the problem of studying the computational complexity of Dempster's rule of combination under certain restrictions on the evidence. A relevant part of the chapter explores how to use a known efficient algorithm to combine hierarchical uncertain evidence to improve the complexity results of combining uncertain evidence in general. In Section 5.2, we present an efficient algorithm to perform the belief operator defined in topological models of evidence. Section 5.3 analyzes the computational complexity of the multi-layer belief model in its most general form. It is based on the computational complexity section of the paper (Pinto Prieto et al., 2023). Finally, Section 5.4 shows a real-life scenario where understanding the peculiarities of the computational complexity of applying Dempster's rule of combination may be advantageous.

Chapter 6: Knowledge Compilation for Combining Uncertain Evidence In this chapter, we show how to compute the result of Dempster's rule of combination (under some restrictions) by constructing a particular propositional logic formula and performing weighted model counting on it with a suitable set of weights. This computation possibly includes the intermediate step of using methods from the field of knowledge compilation to transform the formula into a format that allows efficient application of weighted model counting. We start by introducing weighted model counting and (relevant notions from) knowledge compilation in Section 6.1. Section 6.2 details how to compute this combination rule using weighted model counting. Section 6.3 explains how to apply knowledge compilation to split the whole computation into two phases: an off-line phase that requires exponential time in the worst case and returns a reusable outcome, and an on-line phase that efficiently computes the final result of the rule by using the outcome of the previous step. In Section 6.4, we show how to use this method to efficiently run both phases when the evidence has hierarchical structure. Section 6.5 concludes with a real-life scenario that may benefit from using knowledge compilation to combine uncertain evidence. The content of this chapter is based on an unpublished manuscript co-authored with Ronald de Haan (Pinto Prieto and De Haan, 2024).

Chapter 7: Conclusion This chapter summarizes the main contributions of this thesis and lists some of the research directions that could follow from the presented results. In this chapter, as well as in the previous ones, we assume familiarity with the background introduced in Chapter 2.

Chapter 2

Background

The goal of this thesis is to delve into the problem of combining non-ideal evidence, understanding non-ideal as *uncertain*, *partial*, and *possibly mutually contradictory*. The ultimate objective of combining evidence is to draw robust conclusions from a body of evidence. Drawing these conclusions may involve sound reasoning based on evidence, as may occur in science or justice; result in making decisions, as may happen in autonomous agents or support software; or provide explanations according to the available evidence, as may occur within administrative processes, for example. Therefore, the possibilities for modeling and studying this problem are vast and rooted in many different areas of study. For the specific case of uncertain evidence, we refer to (Halpern, 2017) and (Chaki, 2023). They contain an overview of some consolidated approaches to combine uncertain evidence and a collection of real-life applications for these methods.

In our case, we limit our scope to two concrete approaches, to some extent discussed in the previous books. One is *Dempster-Shafer theory*, a quantitative framework that represents the more basic evidential information as sets of possible states. One of its particularities is that having evidence about a set of possible states A does not provide information either about the subsets of A or about its complement set. For a general overview of the development of this theory, we refer to (Yager and Liu, 2008). The other approach is embedded in formal epistemology. In contrast with the frameworks presented in Halpern's book, we will consider *topological models of evidence*, a topological alternative to the uniform evidence models of (Van Benthem and Pacuit, 2011). Topological spaces are natural models to study evidence and belief, since the conceptual link between them is conveniently captured by topological properties. We will focus on the work of Baltag et al. (2022), that represents evidence both semantically, as sets of possible states, and syntactically in the object language of the proposed logic. For a general overview of this setting, we refer to (Özgün, 2017).

The reason for choosing these two approaches is, in short, that they are independent but complementary, and therefore provide different perspectives on the problem under study. In detail, this translates into the following four points:

1. Both Dempster-Shafer theory and topological models of evidence start by considering a set of possible states S that constitutes the universe of possible elements to be observed.

2. Dempster-Shafer theory models pieces of evidence in terms of subsets of S and adds the notion of uncertainty by allocating some numerical values to these sets.

3. Topological models of evidence also model pieces of evidence as subsets of S, but without numbers or a notion of uncertainty.

4. Both Dempster-Shafer theory and topological models of evidence provide tools for combining evidence with the objective of modeling belief based on it. Dempster-Shafer theory achieves this by aggregating evidence along three dimensions: two quantitative (the number of pieces of evidence and the level of uncertainty) and one qualitative (consistency among pieces of evidence, where two pieces are considered consistent when their set representations have a non-empty intersection). In contrast, topological models of evidence take a purely qualitative approach and do not consider the number of pieces of evidence or the level of uncertainty. However, they elaborate on the qualitative dimension more finely, resulting in a strong notion of belief based on evidence. On a rather abstract level, the notions of belief these theories formalize are linked in the following sense: every proposition believed according to topological models of evidence would also be believed to some non-zero degree in the Dempster-Shafer context, but not necessarily *vice versa*. This distinction arises because the topological models of evidence defined in (Baltag et al., 2022) require consistency not only with respect to the evidence but also with respect to the entire topology generated by this evidence. That is, the formal representation of these models allows to be more demanding with the evidence and distinguish between evidence—or direct observations—and justifications—or reasons to believe in a proposition. From an intuitive point of view, this distinction implies that not every piece of evidence immediately justifies belief; instead, being a piece of justification requires coherence with the body of evidence to which it belongs.

With this theoretical basis in mind, this thesis contributes to the study of combining uncertain, partial, and possibly mutually contradictory evidence in the following way:

- We define a belief model that reproduces topological models of evidence and Dempster-Shafer theory. In other words, we establish a common ground among these approaches, allowing Dempster-Shafer theory to adopt the notion of justification and providing topological models of evidence with a representation of quantitative uncertainty.
- We explore a first attempt to define a qualitative modal logic for this belief model. The goal of this logic is to compare degrees of beliefs and evidential strenght according to the degrees of certainty of the evidence. This enables to logically connect the input and output of the belief model we propose in this work.
- We study the computational complexity of combining evidence by using Dempster-Shafer methods, determining belief through topological models of evidence, and applying our belief model. Although these computations require exponential time in most cases, we identify some alternatives to Dempster-Shafer methods and propose a polynomial-time algorithm for the case of topological models of evidence.
- We propose a concrete algorithmic solution to eliminate the gap between theory and practice in a particular but relevant case of Dempster-Shafer theory.

In the remainder of the chapter, we introduce the terms and notation required to follow the development of all these topics. In Section 2.1, we introduce the relevant concepts of Dempster-Shafer theory for this thesis. In Section 2.2, we define the notions from topological models of evidence that we will adopt. We conclude with Section 2.3, where we establish a common vocabulary to discuss both frameworks simultaneously.

2.1 Dempster-Shafer Theory

Shafer (1976) introduces a theoretical framework that provides tools for modeling evidence that comes from different sources with varying degrees of uncertainty, merging such evidence, and computing the degree of belief from it. This book takes Dempster's rule of combination defined in (Dempster, 1967) as one of its cornerstone, and initiates the so-called *Dempster-Shafer theory*, or *belief function theory*. In this section, we introduce the topic through a list of definitions and some remarks about its strengths and limitations.

Let us assume that there is an agent that collects evidence from several independent sources and aims to compute a degree of belief in a proposition based on the collected evidence. Our challenge is to give a formal representation of all these elements and define formal methods that connect evidence and degrees of belief.

We start by defining the set that will determine the universe of elements relevant to the agent.

Definition 2.1 (Frame of Discernment or Set of Possible States). We call a set S that contains all the elements that may be observed by the sources *frame of discernment* or *set of possible states*.

A set of possible states determines the propositions for which the agent can compute degrees of belief and the potential evidence the agent can collect.

Definition 2.2 (Proposition). We call any subset P of S a *proposition*.

In this context, uncertain evidence is model as *mass functions* defined on the power set of S. The agent will collect a mass function from each source. The subsets of S that receive a positive value through one of these mass functions are the pieces of evidence that the agent counts on. The value given by a mass function represents the *degree of certainty* that the source has about that piece of evidence. Additionally, the value that a mass function gives to the total set S represents the *degree of uncertainty* that the source associates to that evidence input. Since these mass functions are the input received by the agent, they are called *basic belief assignments*.

Definition 2.3 (Basic Belief Assignment). Given a set of possible states S, a *basic belief assignment* over the set S is a function $m : 2^S \to [0,1]$ such that $m(\emptyset) = 0$ and $\sum_{A \subseteq S} m(A) = 1$.

We reserve the term 'basic belief assignment' for those mass functions such that $m(\emptyset) = 0$, but we will also consider others that do not satisfy this property.

Definition 2.4 (Focal Element and Proper Focal Element). Let S be a set of possible states and m be a basic belief assignment. The subsets $A \subseteq S$ for which $m(A) > 0$ are called *focal elements* . In addition, if $A \neq S$ and $m(A) > 0$, we will say that A is a *proper focal element* .

In some parts of this thesis, we will assume specific types of basic belief assignments. For example,

- m is a *non-dogmatic basic belief assignment* if and only if $m(S) > 0$,
- m is a *vacuous basic belief assignment* if and only if $m(S) = 1$,
- m is a *simple basic belief assignment* if and only if there exists $A \subset S$ such that $m(A) > 0$ and $m(S) = 1 - m(A)$, and

- m is a *dichotomous basic belief assignment* if and only if there exists $A \subset S$ such that $m(A) > 0$, $m(S \setminus A) > 0$ and $m(S) = 1 - (m(A) + m(S \setminus A))$.

Note that a simple basic belief assignment has only two focal elements: the total set S and a proper focal element. Similarly, a dichotomous basic belief assignment has three focal elements: the total set S and two proper focal elements A and $S \setminus A$.

Example 2.1 Basic Belief Assignments

Let $S = \{a, b, c\}$ be a set of possible states. Let us assume that an agent receives inputs from three independent sources. Source 1, 2 and 3 respectively return mass functions $m_1 : 2^S \to [0, 1]$, $m_2 : 2^S \to [0, 1]$, $m_3 : 2^S \to [0, 1]$ such that

$$\begin{array}{lll} m_1 : 2^S \to [0,1] & m_2 : 2^S \to [0,1] & m_3 : 2^S \to [0,1] \\ \quad \{a,b\} \mapsto 0.2 & \quad \{b,c\} \mapsto 1 & \quad \{a\} \mapsto 0.2 \\ \quad \{a,b,c\} \mapsto 0.8 & & \quad \{b,c\} \mapsto 0.7 \\ & & \quad \{a,b,c\} \mapsto 0.1 \end{array}$$

Then, Source 1 provides evidence about $\{a, b\}$ with an uncertainty degree of 0.8, Source 2 is fully certain about the input it is providing, and Source 3 is uncertain about its input to degree 0.1. In addition, the agent has evidence about propositions $\{a, b\}$, $\{b, c\}$ and $\{a\}$ with (possibly) different degrees of certainty and from one or more sources.

This set representation of uncertain evidence is key for one of the main theoretical advantages of this theory: modeling *ignorance*. Notice that attending only to Source 1's input, the agent of Example 2.1 does not have information about proposition $\{c\}$, for example. And, according to Source 3, the agent has evidence for $\{b, c\}$ and its complement set $\{a\}$, but the weights associated to them are independent from each other. Stated differently, to have evidence about a proposition, it needs to be observed, independently of other observations. In this case, Source 1 has observed $\{a, b\}$, while Source 3 has observed both $\{b, c\}$ and $\{a\}$—and Source 3 is more certain about the former observation than about the latter. This means that using basic belief assignments to represent evidence not only succeeds in modeling uncertain evidence, but also partial evidence and mutually contradictory evidence. Next definitions clarify how this way of encoding evidence can be used by the agent to obtain degrees of belief.

Definition 2.5 (Belief Function). Given a set of possible states S, a *belief function* is a function $\mathrm{Bel} : 2^S \to [0,1]$ such that

1. $\mathrm{Bel}(\emptyset) = 0$,
2. $\mathrm{Bel}(S) = 1$, and
3. $\mathrm{Bel}\Big(\bigcup_{j=1}^{\ell} A_j\Big) \geq \sum_{\emptyset \neq I \subseteq \{1,\ldots,\ell\}} (-1)^{|I|+1} \mathrm{Bel}\Big(\bigcap_{j\in I} A_j\Big)$,

where $A_j \subseteq S$ for all $j \in \{1, \ldots, \ell\}$.

As observed, the axiomatic definition of a belief function resembles that of probability functions. The only difference is that probability functions assume additivity—i.e., the third condition of Definition 2.5 is an equality in the probabilistic case. In probability theory, an extensively studied alternative to handle uncertainty, a positive degree of belief for the proposition $\{a, b\}$ implies a positive degree of belief for $\{a\}$ or $\{b\}$, due to the additivity axiom. Therefore, in a probabilistic setting, it is not possible to believe in a proposition without believing in some of its singletons. In addition, assuming a total set $S = \{a, b, c\}$, if the degree of belief for $\{a, b\}$ is positive and strictly less than 1, then the agent will believe in $\{c\}$ with positive degree, in particular, with degree $1-$degree of belief in $\{a, b\}$. This means that, in a probabilistic context, having information about one proposition immediately derives information about others. This goes against the notion of ignorance as we have introduced it. Let us see how basic belief assignments and belief functions relate to conclude this discussion.

Definition 2.6 (Belief, Plausibility and Commonality for m). Let S be a set of possible states, $A, B \subseteq S$, and m be a basic belief assignment. Then, we define *belief for m* as

$$\mathrm{bel}_m(B) = \sum_{A \subseteq B} m(A).$$

Belief for m is a belief function (Shafer, 1976, p. 51). Additionally, we define two other relevant functions associated with the notion of belief. We define *plausibility for m* as

$$\mathrm{plau}_m(B) = \sum_{A \cap B \neq \emptyset} m(A)$$

and *commonality for m* as

$$\mathrm{com}_m(B) = \sum_{B \subseteq C} m(C).$$

Note that we skip the subscript m when the associated basic belief assignment is clear from context. More information about these functions can be found in (Shafer, 1976).

Now that we know how to go from basic belief assignments to a belief function, we may say a last word about Example 2.1. In this example, the agent's belief for m_1 is such that $\mathrm{bel}(\{a, b\}) = 0.2$ while $\mathrm{bel}(\{a\}) = \mathrm{bel}(\{b\}) = \mathrm{bel}(\{c\}) = 0$. As we saw, without the relaxation of the additivity axiom in Definition 2.5 this would not be possible. This characteristic is particularly useful in cases where partial evidence plays an important role—for instance, in some medical diagnosis scenarios (Verbert et al., 2017).

Up to here, we have introduced one of the strengths of this theory that makes it relevant for us. We summarize it in the following paragraph.

Dempster-Shafer theory models uncertain, partial and potentially mutually contradictory evidence through basic belief assignments m. From this evidence encoded in m, this theory defines a function that returns degrees of belief for any proposition.

The second part of this introduction to Dempster-Shafer theory focuses on how to combine different sources of evidence, i.e., how to combine a collection of basic belief assignments $m_1, \dots, m_\ell$ to obtain a belief function based on all of them. We will introduce three different rules of combination. Let us start with the one introduced by Dempster (1967).

Definition 2.7 (Dempster's Rule of Combination or Normalized Conjunctive Rule). Let m_1 and m_2 be basic belief assignments over the same set S of possible states and $A_1, \dots, A_r$ and $B_1, \dots, B_s$ all subsets of S such that $m_1(A_j) > 0$ and $m_2(B_k) > 0$, respectively. Moreover, suppose that $\sum_{A_j \cap B_k = \emptyset} m_1(A_j)m_2(B_k) < 1$. Then the following basic belief assignment m, also denoted by $m_1 \oplus m_2$, is the result of applying *Dempster's rule of combination* to m_1 and m_2:

$$m_1 \oplus m_2(C) = \begin{cases} 0 & \text{if } C = \emptyset, \\ \frac{\sum_{A_j \cap B_k = C} m_1(A_j)m_2(B_k)}{K} & \text{otherwise} \end{cases}$$

where K is the normalization factor $1 - \sum_{A_j \cap B_k = \emptyset} m_1(A_j)m_2(B_k)$.

In some cases, we call this rule *normalized rule* for abbreviation, and the resulting basic belief assignment *combined mass function* for differentiation. This rule provides the second strength of this theory that is of special interest to us. We summarize it in the following sentence.

Dempster-Shafer theory provides combination rules that allow to merge uncertain, partial and possibly mutually contradictory evidence from independent sources, obtaining a combined body of evidence that may be used to obtain degrees of belief for propositions.

Although we have defined belief functions in general, the most applied side of this theory focuses on a special version of them, the so-called *support functions*.

Definition 2.8 (Support Function). Let S be a set of possible states and $\mathrm{bel} : 2^S \to [0,1]$ be a belief function. If there are some basic belief assignments $m_1, \dots, m_\ell$ such that bel is a belief function for $(m_1 \oplus \dots \oplus m_\ell)$, then bel is a *support function*.

If all the basic belief assignments $m_1, \dots, m_\ell$ are simple belief assignments—i.e., they each have a unique proper focal element—bel is said to be a *separable support function*.

Dempster's rule of combination performs in a satisfactory way in many situation, as we can see in the already mentioned sources. Nevertheless, there are some limitations recurrently remarked:

1. The straightforward implementation of this method takes exponential time in general (Barnett, 1981; Orponen, 1990). We explore this issue in Chapters 5 and 6.
2. The normalization factor of this rule has been controversial, with arguments for and against in the literature (Haenni, 2005; Lefevre et al., 2002; Murphy, 2000; Pearl, 1990; Sentz and Ferson, 2002; Smets, 2007; Zadeh, 1986). All these sources agree that highly conflicting evidence—those situations where there are many focal elements from different basic belief assignments that do not intersect with each other—deserves special attention. We will introduce two other rules whose definitions are partially motivated by this discussion.
3. Dempster's rule of combination is defined for independent sources of evidence, which limits its application in many real-life scenarios. Denœux (2008) analyzes this limitation and proposes a new family of combination rules that overcome this problem. Cases of dependent sources of evidence are out of the scope of this thesis.

The first alternative to Dempster's rule of combination is defined in (Yager, 1987) and received further support in (Smets and Kennes, 1994). This belief model has been further justified in (Pichon and Denoeux, 2010). One characteristic of this belief model is that it does not require the basic belief assignments m to assign 0 to the empty set. Actually, the value assigned to the empty set can be interpreted as an estimation of the amount of conflict (Destercke and Burger, 2013). For the sake of clarity, we will refer to this particular case of basic belief assignments as *general mass functions*.

Definition 2.9 (General Mass Functions). Given a set of possible states S, a *general mass function* over the set S is a function $m : 2^S \to [0,1]$ such that $\sum_{A \subseteq S} m(A) = 1$.

Definition 2.10 (Belief for a General Mass Function). Let S be a set of possible states, $A, B \subseteq S$, and m be a general mass function. Then, we define *belief for* m as

$$\mathrm{bel}(B) = \begin{cases} \left(\sum_{A \subseteq B} m(A)\right) - m(\emptyset) & \text{if } B \neq S, \\ \sum_{A \subseteq B} m(A) & \text{otherwise.} \end{cases}$$

Definition 2.11 (Transferable Rule of Combination or Unnormalized Conjunctive Rule of Combination). Let m_1 and m_2 be general mass functions over the same set S of possible states. Let $A_1, \ldots, A_r$ and $B_1, \ldots, B_s$ be all subsets of S such that $m_1(A_j) > 0$ and $m_2(B_k) > 0$, respectively. Then the following general mass function m, also denoted by $m_1 \boxplus m_2$, is the result of applying *transferable rule of combination* to m_1 and m_2:

$$(m_1 \boxplus m_2)(C) = \sum_{A \cap B = C} m_1(A) \cdot m_2(B).$$

Sometimes we will refer to this rule as *unnormalized rule of combination* for brevity.

The second alternative for Dempster's rule of combination that we will cover is known as *unnormalized disjunctive rule of combination* or *disjunctive rule of combination* for brevity. It was defined in (Dubois and Prade, 1986) and reinforced in (Dubois and Prade, 1992; Smets, 1993). In this case as well, it is allowed for the basic belief assignments to assign a positive value to the empty set, so we will refer to them as general mass functions. As we can see in the following definition, the main difference compared to the previous rules is that we combine mass functions according to the union of focal elements instead of their intersection. The motivation for this change relies on the reliability of the sources: when using the intersection we are assuming that all the sources are reliable, while choosing the union implies that at least one source is reliable but we do not know which one (Denœux, 2008).

Definition 2.12 (Unnormalized Disjunctive Rule of Combination). Let m_1 and m_2 be general mass functions over the same set S of possible states and $A_1, \ldots, A_r$ and $B_1, \ldots, B_s$ all subsets of S such that $m_1(A_j) > 0$ and $m_2(B_k) > 0$, respectively. Then the following general mass function m, denoted by $m_1 \circledcirc m_2$

as well, is the result of applying *unnormalized disjunctive rule of combination* to m_1 and m_2:

$$(m_1 \circledcirc m_2)(C) = \sum_{A \cup B = C} m_1(A) \cdot m_2(B).$$

Our (non-exhaustive) introduction to the Dempster-Shafer theory ends here. We refer to the references provided in Chapter 1 and this section for further information on this.

2.2 Topological Models of Evidence

Formal epistemology is a field of study that uses formal methods and tools from mathematics and logic to study notions and issues of epistemological interest. The main epistemological notions it investigates and that are of particular interest to this dissertation are notions of evidence, justification and belief. Hintikka (1962) defined a first epistemic modal logic for reasoning about these concepts using standard possible states semantics based on Kripke structures. Van Benthem and Pacuit (2011) introduce the notion of evidence as information that an agent may gather but is potentially false or misleading, moving away from the definition of evidence as *factive evidence*. This new interpretation of evidence opens the door to considering more realistic evidence, such as partial or possibly mutually contradictory evidence. The evidence models defined by Van Benthem and Pacuit (2011) follows neighborhood semantics instead of relational semantics. This is because neighborhood semantics allows the weakening of the assumptions about the agent's knowledge and belief and explicit representations of evidence, making it a more flexible semantics for this context. We refer to (Pacuit, 2017) for a detailed justification.

From the work of Van Benthem and Pacuit (2011), we are especially interested in the uniform evidence model, which assigns the same set of pieces of evidence to every possible state. In their words, this corresponds to working with agents who are evidence-introspective, or agents who are fully aware of the available evidence and its implications. These evidence models are reformulated in (Baltag et al., 2022) by changing neighborhood semantics to topological semantics and modifying their definition of belief. These changes kept both alternatives equivalent in the finite case and brought some technical improvements in the infinite one (see Example 1, Example 2 and Proposition 1 of (Baltag et al., 2022, Section 4)). While these improvements are not relevant for our particular case, the combination of their definition of belief and topological semantics offers an explicit connection among belief, a notion of justification and topological properties of the evidence. In Section 2.3, we will see how this explicitness is particularly useful for connecting epistemic evidence models and Dempster-Shafer theory. In the remainder of this section, we will list the technical concepts from (Baltag et al., 2022) required to follow the rest of the thesis.

Let us assume that an agent possesses a collection of pieces of evidence and aims to decide whether they believe in a certain proposition according to that evidence, ensuring that their belief will be consistent regardless of the gathered evidence. Our challenge here is to formalize these concepts and define an operator for belief that meets this constraint. Note that we do not consider our evidence to be uncertain in this case.

In this context, a *set of possible states* and a *proposition* follow definitions 2.1 and 2.2 as well. Additionally, a piece of evidence is represented as a subset of S. Under our assumptions, the agent possesses a collection of pieces of evidence named *basic* or *direct evidence*. This basic evidence may come from direct observation, measurements, testimony from others, and similar sources. They are basic pieces of evidence in this sense. Notice that the gathered pieces of evidence are not supposed to be independent, nor are their sources.

Definition 2.13 (Basic Evidence Frame). Let S be a set of possible states and $\boldsymbol{\mathcal{E}}$ be a non-empty subset of 2^S such that $\emptyset \notin \boldsymbol{\mathcal{E}}$ and $S \in \boldsymbol{\mathcal{E}}$. We call the tuple $(S, \boldsymbol{\mathcal{E}})$ a *basic evidence frame*, and $\boldsymbol{\mathcal{E}}$ *set of basic pieces of evidence* or *basic evidence set*.

Sometimes we refer to the elements of $\boldsymbol{\mathcal{E}}$ simply as *pieces of evidence* for short. Notice that the constraint $S \in \boldsymbol{\mathcal{E}}$ ensures that tautologies are always evidence, and $\emptyset \notin \boldsymbol{\mathcal{E}}$ that contradictions are never evidence. The definition of belief that we introduce later will clarify these restrictions.

From a basic evidence frame $(S, \boldsymbol{\mathcal{E}})$, an agent may derive some information that, in this context, is known as *combined evidence* and *evidence-based arguments*. We will introduce the formal definitions of these concepts through the notion of *topological space*.

Definition 2.14 (Topological Space and Open Sets). A *topological space* is a pair $(S, \boldsymbol{\tau})$, where S is a non-empty set and $\boldsymbol{\tau}$ is a family of subsets of S such that $S, \emptyset \in \boldsymbol{\tau}$, and $\boldsymbol{\tau}$ is closed under finite intersections and arbitrary unions. Given a topological space $(S, \boldsymbol{\tau})$, the set S is called a *space* and the family $\boldsymbol{\tau}$ is called a *topology* on S. In addition, the elements of $\boldsymbol{\tau}$ are called *open sets* or *opens* in the space.

Onwards, we will identify our pieces of evidence as open sets of a topology. The treatment of open sets as pieces of evidence dates back to (Troelstra and Van Dalen, 1988) and is adopted from topological semantics for intuitionistic logic. It moreover has applications in domain theory (Vickers, 1989) and formal learning theory (Kelly, 1996). Next definitions explains how to technically connect evidence and open sets. For further details on epistemic interpretations of topological spaces, we refer to (Özgün, 2017).

Definition 2.15 (Subbasis, Basis and Generated Topology). A topology $\boldsymbol{\tau}$ on a space S can also be generated from an arbitrary subset of 2^S. Given any family $\boldsymbol{\mathcal{E}} \subseteq 2^S$ of subsets of S, there exists a unique, smallest topology $\boldsymbol{\tau_{\mathcal{E}}}$ with $\boldsymbol{\mathcal{E}} \subseteq \boldsymbol{\tau_{\mathcal{E}}}$ (Dugundji, 1965, Theorem 3.1, p. 65). The family $\boldsymbol{\tau_{\mathcal{E}}}$ consists of $\emptyset$, S, all finite intersections of $\boldsymbol{\mathcal{E}}$, and all arbitrary unions of these finite intersections. $\boldsymbol{\mathcal{E}}$ is called a *subbasis* for $\boldsymbol{\tau_{\mathcal{E}}}$ and $\boldsymbol{\tau_{\mathcal{E}}}$ is said to be *generated* by $\boldsymbol{\mathcal{E}}$. We say that the closure under finite intersections of $\boldsymbol{\mathcal{E}}$ forms a *topological basis* for $\boldsymbol{\tau_{\mathcal{E}}}$.

Definition 2.16 (Combined Evidence). Considering a basic evidence frame $(S, \boldsymbol{\mathcal{E}})$, an element of the topological basis for $\boldsymbol{\tau_{\mathcal{E}}}$ generated by $\boldsymbol{\mathcal{E}}$ is called *combined evidence*. In other words, given a basic evidence frame $(S, \boldsymbol{\mathcal{E}})$, the set of combined evidence is the set of finite intersections of members of $\boldsymbol{\mathcal{E}}$.

Definition 2.17 (Evidential Topology). Considering a basic evidence frame $(S, \boldsymbol{\mathcal{E}})$, the topology $\boldsymbol{\tau_{\mathcal{E}}}$ generated by $\boldsymbol{\mathcal{E}}$ is called the *evidential topology*.

Definition 2.18 (Evidential Support). Given a basic evidence frame $(S, \boldsymbol{\mathcal{E}})$, a proposition $P \subseteq S$, and a piece of evidence $E \in \tau_{\boldsymbol{\mathcal{E}}}$, E *supports* P if and only if $E \subseteq P$. Alternatively, we will say E is *evidential support* for P if and only if $E \subseteq P$.

Definition 2.19 (Evidence-Based Argument). Given a basic evidence frame $(S, \boldsymbol{\mathcal{E}})$ and a proposition $P \subseteq S$, an *evidence-based argument for* P is an element $T \in \boldsymbol{\tau_{\mathcal{E}}} \setminus \{\emptyset\}$ such that $T \subseteq P$. Topologically, an evidence-based argument is just a non-empty open set in a topology. We may say *argument* for short.

Topological models of evidence represent partial and mutually contradictory evidence and can combine it to form a richer set of evidence-based information.

One of the goals of this approach is to define consistent belief based on possibly mutually contradictory and partial evidence. For topological models of evidence, Baltag et al. (2022) define belief by using the notion of *denseness*, a topological notion closely related to consistency.

Definition 2.20 (Denseness). Let $(S, \boldsymbol{\tau})$ be a topological space. A set $P \subseteq S$ is called *dense* in S (with respect to topology $\boldsymbol{\tau}$) if and only if $P \cap T \neq \emptyset$ for all

$T \in \boldsymbol{\tau}$ such that $T \neq \emptyset$. When it is contextually clear, we call a set *dense* only and avoid mention of the relevant space and its respective topology.

Given $(S, \boldsymbol{\mathcal{E}})$ a basic evidence frame and $\boldsymbol{\tau_{\mathcal{E}}}$ the corresponding evidential topology, we sometimes say that $P \subseteq S$ *consistent with* $\boldsymbol{\tau_{\mathcal{E}}}$ when P is a dense set with respect to $\boldsymbol{\tau_{\mathcal{E}}}$.

Definition 2.21 (Mutual consistency). Let $(S, \boldsymbol{\mathcal{E}})$ be a basic evidence frame and $\boldsymbol{\tau_{\mathcal{E}}}$ be the corresponding evidential topology. We say that two pieces of evidence $E, E' \in \boldsymbol{\tau_{\mathcal{E}}}$ are *mutually consistent* if $E \cap E' \neq \emptyset$, and mutually inconsistent otherwise.

Notice that $\boldsymbol{\tau_{\mathcal{E}}}$ (as well as $\boldsymbol{\mathcal{E}}$) can host mutually inconsistent pieces of evidence, since we did not impose any constraints on $\boldsymbol{\mathcal{E}}$ to eliminate such cases.

Baltag et al. (2022) use this property of being dense in the set of possible states with respect to a topology to establish their definition of *evidence-based justification*. Justifications will be another fine-grained representation in topological models of evidence to differentiate among various kinds of evidence, just as basic evidence, combined evidence and arguments are.

Definition 2.22 (Evidence-Based Justification). Let $(S, \boldsymbol{\mathcal{E}})$ be a basic evidence frame and $\boldsymbol{\tau_{\mathcal{E}}}$ be the topology generated by $\boldsymbol{\mathcal{E}}$. An *evidence-based justification* or *justification* is a dense open subset of $\boldsymbol{\tau_{\mathcal{E}}}$. A *justification for* P is an element $J \in \boldsymbol{\tau_{\mathcal{E}}} \setminus \{\emptyset\}$ such that J is dense in S with respect to $\boldsymbol{\tau_{\mathcal{E}}}$ and $J \subseteq P$.

Therefore, a justification is an argument that is consistent with any other argument based on the available evidence. This property entails that the agent always has consistent beliefs, even when the belief is formed based on a set of possibly mutually inconsistent pieces of evidence (Baltag et al., 2022). In Definition 2.23 we present two equivalent definitions of belief according to (Baltag et al., 2022, Proposition 2).

Definition 2.23 (Justified Belief in Topological Models of Evidence). Let $(S, \boldsymbol{\mathcal{E}})$ be a basic evidence frame and $\boldsymbol{\tau_{\mathcal{E}}}$ be the topology generated by $\boldsymbol{\mathcal{E}}$. A proposition $P \subseteq S$ is *justified believed* if and only if there is a justification for P in $\boldsymbol{\tau_{\mathcal{E}}}$. Equivalently, a proposition $P \subseteq S$ is *justified believed* if and only if P includes some dense open set.

Example 2.2 From Basic Evidence to Justification

Let $S = \{a, b, c, d, e\}$ be a set of possible states and $\mathcal{E} = \big\{\{a, b\}, \{b, c, d\}, \{d, e\}\big\}$ be a basic evidence set (the evidence directly gathered by the agent). The following diagrams show the different sets that we can obtain from $\mathcal{E}$ according to the previous definitions.

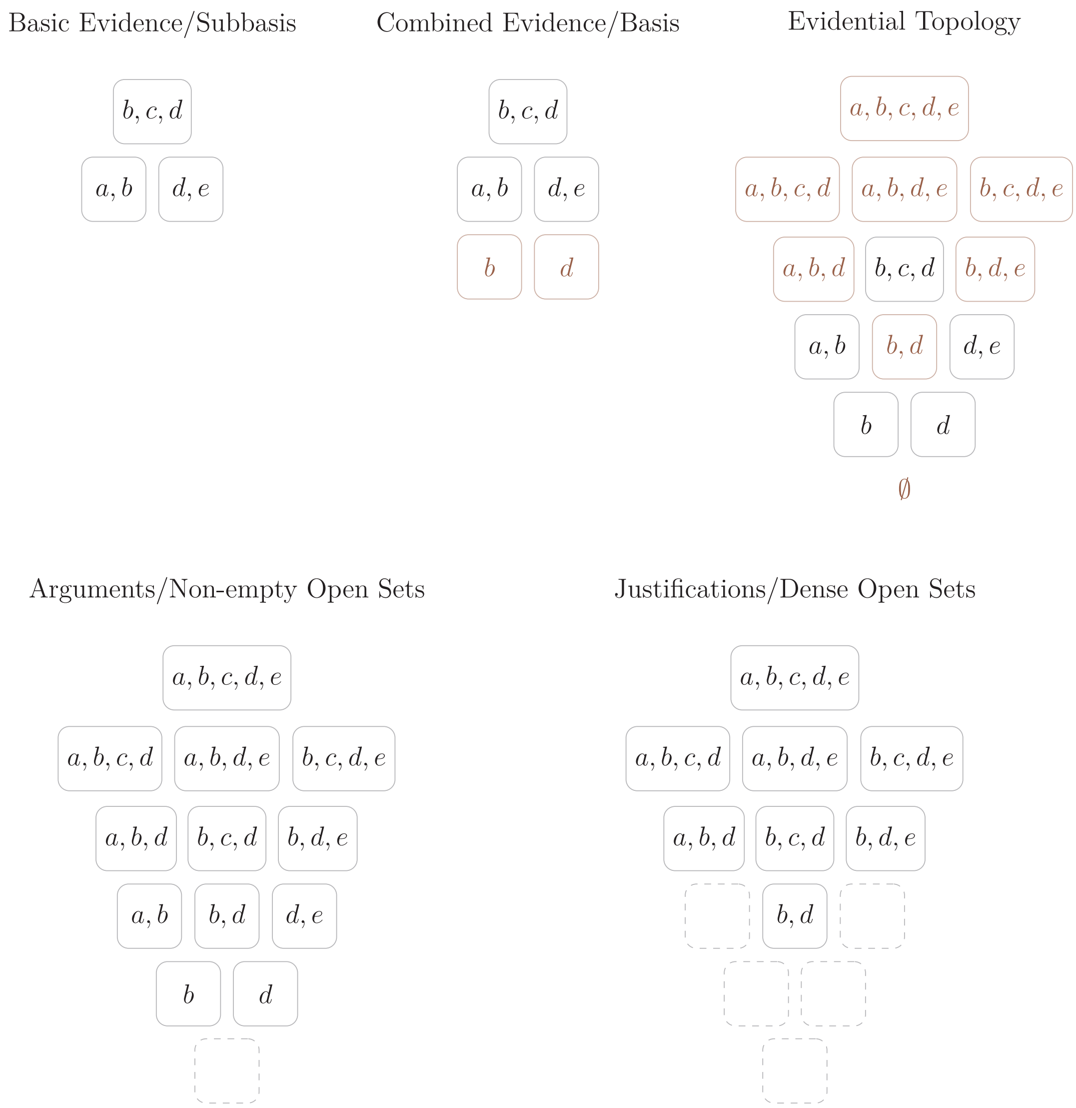


Figure 2.1: Visual representation of the evidence-based sets defined in this section. Each box represents a set formed by the elements enclosed in it. Following the reading order, colored sets indicate that they were not included in the previous structure. Pointed squares indicate that the current structure is missing the corresponding set with respect to the previous one.

As observed in the previous example, not every piece of evidence or argument constitutes a piece of justification for agents' beliefs. Underlying intuition here is that an agent have high demands for what constitutes a justification for their beliefs. In particular, observing something alone might not be good enough, the agent also requires consistency with all the other observations they have made.

Topological models of evidence define consistent belief based on possibly partial and contradictory evidence, due to the explicit differentiation between evidence and justification.

From (Baltag et al., 2022; Özgün, 2017) and this brief introduction to topological models of evidence, we can identify two limitations of this approach as a tool for studying the problem addressed in this thesis:

1. Due to the purely qualitative nature of these models, it is not possible to represent varying degree of certainty for each piece of evidence within the model. The results in Chapters 3 and 4 serve as a first step towards a quantitative extension of these epistemic models.
2. A computational complexity analysis of these particular models, as well as an algorithmic method for combining evidence according to them in practice, is missing in the literature. Section 5.2 shows partial results in this direction.

2.3 A Common Ground

The preceding sections present Dempster-Shafer theory and topological models of evidence as interesting frameworks to bring light to the problem of *combining uncertain, partial and possibly mutually contradictory evidence*. The first challenge we will address is to combine these approaches, aiming to provide a common (and hopefully richer) framework. To this end, we first need a common vocabulary and notation. We dedicate this section to this purpose, as well as to highlighting their compatibility and complementary nature.

Firstly, we should note that the strategies employed by Dempster-Shafer theory and topological models of evidence are notably similar:

1. They allow the agent to gather non-ideal evidence, understanding non-ideal as uncertain, partial and possibly mutually contradictory evidence.
2. They provide us with a method for combining the available evidence and get some extra information according to some features of that evidence.
3. They define a notion of evidence-based belief for the agent.

Secondly, they use compatible representations of evidence. Specifically, they both start with a set of possible states S and define pieces of evidence as subsets of S. In both cases, the most basic information relies on these sets, not on the singletons of S. Throughout this thesis, we assume that a set of possible states is *finite*, reflecting our ultimate motivation of applying these constructs to real-life scenarios.

Finally, Dempster-Shafer theory and topological models of evidence also differ in some aspects. However, these distinctions are not inconvenient but rather make them complementary. For instance, Dempster-Shafer theory accepts uncertain, possibly partial and mutually contradictory evidence from independent sources. Topological models of evidence, in contrast, accept possibly partial and mutually contradictory evidence that may come from dependent sources. In our particular case, we will bring the topological framework closer to Dempster-Shafer theory by assuming independence between sources and defining quantitative evidence frames. We will use these structures to represent uncertain evidence.

Definition 2.24 (Quantitative and Qualitative Evidence Frames). Let S be a set of possible states, $\boldsymbol{\mathcal{E}}$ be a non-empty subset of 2^S such that $\emptyset \notin \boldsymbol{\mathcal{E}}$ and $S \notin \boldsymbol{\mathcal{E}}$, and $\boldsymbol{\mathcal{E}}^Q$ be a nonempty subset of $\boldsymbol{\mathcal{E}} \times (0,1)$ containing exactly one element for each element of $\boldsymbol{\mathcal{E}}$. We call the tuple $(S, \boldsymbol{\mathcal{E}}^Q)$ a *quantitative evidence frame,* and $\boldsymbol{\mathcal{E}}^Q$ a *quantitative evidence set.* Additionally, we will say that $\boldsymbol{\mathcal{E}}$ is a *qualitative evidence set,* and $(S, \boldsymbol{\mathcal{E}})$ is a *qualitative evidence frame.* Note that in topological models of evidence $(S, \boldsymbol{\mathcal{E}})$ is called a basic evidence frame. We call it a qualitative evidence frame from now on to emphasize the fact that this is part of a quantitative frame without the quantitative component Q.

As before, we refer to $P \subseteq S$ as a proposition, and we introduce the name *propositional content* to specifically denote elements of $\boldsymbol{\mathcal{E}}$.

Definition 2.25 (Propositional Content). By *propositional content* of a piece of evidence, we mean the information provided by a piece of evidence without regard to how uncertain that piece of evidence is.

For any element $(E, p) \in \boldsymbol{\mathcal{E}}^Q$, E represents the propositional content of the evidence and p is its degree of certainty. Given a pair $(E, p) \in \boldsymbol{\mathcal{E}}^Q$, the value $1 - p$ represents the uncertainty of the given piece of evidence (and not the certainty of $S \setminus E$). In this particular setting, we sometimes use “degree of certainty” and “degree of uncertainty” of a piece of evidence interchangeably in our conceptual explanations since the latter is fully determined by the value of the former p—once we have one of the values, we have the other. In Dempster-Shafer terminology, the set $\boldsymbol{\mathcal{E}}^Q$ can be interpreted as a set of non-dogmatic (nor vacuous) simple basic belief assignments. Each element $(E, p) \in \boldsymbol{\mathcal{E}}^Q$ represents a non-dogmatic simple basic belief assignment m such that $m(E) = p$. We choose the open interval $(0, 1)$ to avoid considering exceptions in the following sections. However, this assumption is not a significant limitation since the aim of the approach is to model uncertainty. Note that this representation of the evidence does not include the case of non-simple basic belief assignments—i.e., basic belief assignments with more than one proper focal element.

We will use the term *piece of evidence* to refer to an element of $\boldsymbol{\mathcal{E}}$ or an element of $\boldsymbol{\mathcal{E}}^Q$ depending on the context. Notice that, contrary to the definition of basic evidence frame of topological models of evidence, the total set S is not included in $\boldsymbol{\mathcal{E}}$. We impose $S \notin \boldsymbol{\mathcal{E}}$ because, following Dempster-Shafer theory, the mass values assigned to S will represent degrees of uncertainties about evidence, not the degree of certainty in S (which is always 1). Another difference with respect to the topological models of evidence of Baltag et al. (2022) is that we use the term 'evidence' for basic pieces of evidence, while they use this word for what they call combined evidence. A last note on terminology, when we stick to Dempster-Shafer notation, we use the term 'evidence body' or 'body of evidence' to refer to the collection of basic belief assignments. Baltag et al. (2022) use the words 'body of evidence' for a concrete formal structure that is not contemplated in this thesis.

There are also differences in the combination method used by each approach. Dempster-Shafer theory employs arithmetic rules that take into account three characteristics of the evidence: the number of pieces of evidence (sum), the degree of uncertainty of each piece of evidence (product), and the consistency among pieces of evidence (intersection or union). In contrast, topological model of evidence only consider the consistency among pieces of evidence. Besides these differences, these approaches are compatible. On the one hand, topological models of evidence can be enriched by defining a quantitative version including the other two features considered by Dempster-Shafer combination rules. On the other hand, when applying the combination rules defined in Section 2.1, the focal elements of the combined mass functions are always within the evidential topology corresponding to the quantitative evidence set formed by the focal elements of the basic belief assignments to combine.

Lastly, the definition and understanding of evidence-based belief differs in each case, yet remains compatible. For example, considering dense open sets in Dempster-Shafer theory may enhance robustness in cases of highly conflicting evidence. Another benefit of combining both notions of belief is bringing topological models of evidence closer to practical applications.

For the remaining definitions given in Section 2.2, we apply them to the quantitative evidence frame in general, except for 'justifications'. We will relax their definition to be special kinds of arguments that the agent uses to support their beliefs and the form of such justifications depends on the agent's evidential demands. Given a quantitative evidence frame $(S, \boldsymbol{\mathcal{E}}^Q)$, our ultimate goal is to define a belief function in the sense of Definition 2.5 which admits as input not only a body of possibly mutually inconsistent and uncertain evidence $\boldsymbol{\mathcal{E}}^Q$ but also the evidential demands of the agent. We elaborate more on this in Chapter 3. Additionally, we will also use the term 'mass function' to talk about mass functions defined not necessarily for elements of 2^S; in particular, we will define a mass function for elements of $2^{\boldsymbol{\mathcal{E}}}$. Therefore, we keep in mind a more general definition of mass functions.

Definition 2.26 (Mass Function Over a Set). Given a nonempty set X, a *mass function* over the set X is a function $m : 2^X \to [0, 1]$ such that $\sum_{A \subseteq X} m(A) = 1$.

Throughout this dissertation we use standard order-theoretic notions such as total pre-order, minimal/maximal elements, minimum/maximum elements in an order, etc. For all these definition we refer to (Roman, 2008).

Notational Conventions

Aiming to help the reader to follow the text smoothly, we have set a notation code. Lowercase letters refer to possible states. Uppercase letters are used to specify sets of possible states. When using the notation introduced in this section, we will use E to represent basic pieces of evidence, T to represent the elements of a topology, S to represent the set of all the possible states, and P to represent a proposition. Sets of the previous sets are named in bold capital letters. Some examples are $\boldsymbol{E}$ to specify any set of pieces of evidence and $\boldsymbol{\mathcal{E}}$ to specify the set of all the pieces of evidence. Finally, the subsets of $2^{\boldsymbol{\mathcal{E}}}$ are denoted by blackboard bold capital letters (such as $\mathbb{M}$).

In Section 5.1 and Chapter 6 we will follow Dempster-Shafer notation, since the results of those parts strictly belong to Dempster-Shafer framework. In those cases, we will use F to talk about focal elements (including S when it applies) and A, B or C to represent proper focal elements. In this context, we will name sets of sets of possible states with a calligraphic capital letter. For instance, $\mathcal{F}$ is used to name a set of focal elements. We also set a code for the subindexes. The letter n is reserved to indicate the number of elements of S and the subindex i to talk about numbers between 1 and n. The letter ℓ is reserved to indicate the number of pieces of evidence or the number of basic belief assignments. We will use the subindexes j and k to refer to numbers between 1 and ℓ. We conclude the section with a clarification table of terms.

This Thesis	**Dempster-Shafer Theory**	**Topological Models of Evidence**
Uncertain, *partial* and possibly *mutually contradictory* evidence	*Uncertain*, *partial* and possibly *mutually contradictory* evidence	*Partial* and possibly *mutually contradictory* evidence
Independent sources	*Independent* sources	Possibly *dependent* sources
Mutually contradictory $\sim$ inconsistent $\sim$ conflicting $\sim$ empty intersection		
Set of possible states $S = \{s_1, \ldots, s_i, \ldots, s_n\}$	Frame of discernment	Set of possible states
Proposition $P \subseteq S$	Proposition	Proposition
Quantitative evidence set $\boldsymbol{\mathcal{E}}^Q = \Big\{(E_1, p1), \ldots, (E_j, p_j), \ldots, (E_\ell, p_\ell)\Big\}$	Non-dogmatic simple basic belief assignments $m_1, \ldots, m_j, \ldots, m_\ell$	
Qualitative evidence set $\boldsymbol{\mathcal{E}} = \{E_1, \ldots, E_j, \ldots, E_\ell\}$	Proper focal elements of simple basic belief assignments $A_1, \ldots, A_j, \ldots, A_\ell$	Basic evidence set without S
Evidence	Body of evidence	Basic evidence
Quantitative piece of evidence $(E_j, p_j) \in \boldsymbol{\mathcal{E}}^Q$	Proper focal element of m and its value through m	
Uncertainty degree of $(E_j, p_j) \in \boldsymbol{\mathcal{E}}^Q$: $1 - p_j$	If m simple basic belief assignment with proper focal element E_j, $m(S)$	
Propositional content E_j from $(E_j, p_j) \in \boldsymbol{\mathcal{E}}^Q$	Focal element of m	Basic piece of evidence
Mass function over S such that $m(\emptyset) = 0$	Basic belief assignment	
To be defined in Chapter 3	Combined mass function through conjunctive combination rules	Combined evidence set
Topology generated by $\boldsymbol{\mathcal{E}}$		Evidential topology
Argument		Argument
Justification		Justification is one particular example
Degree of belief	Support function	Justified belief modal operator

Table 2.1: Comparison of terms used in this thesis, Dempster-Shafer theory, and topological models of evidence.

Chapter 3

Multi-Layer Belief Model for Combining Uncertain Evidence

In the previous chapter, we explored several models that are used in the literature to merge evidence and generate belief based on it. Each of these models focuses on one feature of the evidence—uncertainty, inconsistency, etc.—and introduces precise concepts and tools to successfully manage it. Therefore, having a single tool kit to represent them all would undoubtedly be useful. That is the goal of this chapter: defining a *multi-layer* framework that includes the different components of the previously mentioned belief models and, consequently, that allows us to switch from one to the other with the simplicity of changing parameters. To this end, we will study separately the qualitative and the quantitative components of these belief models. We will define a method that treats these components individually and, afterward, puts them together. This explicit division facilitates a further understanding of the represented belief models since we can see which component is responsible for certain behaviors—the qualitative one, the quantitative one, or the merging process. Overall, the multi-layer belief model can be easily adapted to the situation, and, as we will see in Chapter 5, without an extra cost in the computational complexity.

The remaining of this chapter is organized as follows. In Section 3.1, we introduce the technical definition of the model consisting in three parts: the qualitative layer, the quantitative layer, and the bridging layer. In Section 3.2, we explore the mathematical properties of the model and we compare it with other well-known belief models. In particular, we show that the multi-layer belief model manages to represent the belief functions obtained by Dempster's rule of combination and the transferable belief model, as well as, to recover the notion of binary belief represented in the topological models of evidence. In addition, we outline how this model can provide interesting alternatives to the disjunctive rule of combination and the transferable belief model. In Section 3.3, we end the chapter by showing an illustrative example to understand the value and the potential use of this model.

3.1 The Model

In Chapter 2, we introduced the works that inspire the definition introduced in this section, with (Shafer, 1976) and (Baltag et al., 2022) as cornerstones. With the multi-layer belief model we aim to propose a method for measuring degrees of beliefs for an agent that possesses possibly *mutually contradictory*, *partial*, and *uncertain* pieces of evidence. We will do that by enriching our framework with enough expressive power to represent *ignorance* and *degrees of uncertainty*, and to differentiate among *basic evidence*, *combined evidence*, *argument*, and *justification* according to Section 2.2. As a result, we will get a belief model adaptable to the *evidential demands* of the agent and their *context*. We will detail what this means later on.

As its name suggests, this belief model is built on three different layers that will interact among them at the end of the process. The first layer, called the *qualitative layer*, works with the propositional content of the basic pieces of evidence—i.e., elements of $\boldsymbol{\mathcal{E}}$—and identifies a set of justifications that represents the agent's evidential demands. This layer is inspired by the topological models of evidence defined in Section 2.2. The second layer, called the *quantitative layer*, focuses on combining the degrees of uncertainty about evidence. This layer is inspired by the combination rules discussed in Section 2.1, but in this case, the combined mass function is defined for sets of pieces of evidence, i.e., its domain is $2^{\boldsymbol{\mathcal{E}}}$. Since our goal is to compute degrees of belief based on justifications—which are sets of possible states—and the combined uncertainty values are assigned to sets of sets of possible states instead, we need to transfer these values from the sets of sets of possible states to sets of possible states. To this end, we define a last layer, called the *bridging layer*, that connects the values of the mass function obtained in the second layer to the justifications of the first one. As a result, we obtain a belief function which is able to compute degrees of belief according to the evidential demands of the agent and based on a possibly mutually contradictory, partial, and uncertain body of evidence.

3.1.1 Qualitative Layer

The aim of this layer is to generate a set of justifications based on the available evidence and the agent's evidential demands. Following Definition 2.19, given a qualitative evidence frame $(S, \boldsymbol{\mathcal{E}})$, the set $\boldsymbol{\tau}_{\boldsymbol{\mathcal{E}}} \backslash \{\emptyset\}$ represents the set of *arguments* available to the agent. Any argument is a potential piece of justification for belief. What will determine whether an argument is a justification are the conditions that the agent requires for it to be convincing enough to base belief on. These requirements that the agent imposes on the evidence are what we call *evidential demands*. They are independent of the proposition for which the agent wants to compute degree of belief and the uncertainty values the evidence may have. The more restrictive these conditions are, the more cautious the agent is. For example, a doctor would be more cautious evaluating diagnostic tests when prescribing a medication with risky side effects than when prescribing low-risk medication. Note that evidential demand refers to a constraint given by agents independently of the sources of evidence. It should not be confused with agents' degrees of trust on the sources or with agents' prior beliefs.

To distinguish arguments from justifications and formalize an agent's evidential demands, we use the notion of *frame of justification*. Given $(S, \boldsymbol{\mathcal{E}})$, a frame of justification $\boldsymbol{\mathcal{J}}$ is a subset of $\boldsymbol{\tau_{\mathcal{E}}} \setminus \{\emptyset\}$. Depending on the modelled situation, one can think of natural constraints to impose on frames of justifications. For example, we can think about frames of justification inspired by Dempster-Shafer theory and the framework of topological models of evidence. Given a qualitative evidence frame $(S, \boldsymbol{\mathcal{E}})$:

1. The *Dempster-Shafer frame of justification* , denoted by $\boldsymbol{\mathcal{J}}^{DS}$, is the set of all arguments, that is, $\boldsymbol{\mathcal{J}}^{DS} = \boldsymbol{\tau_{\mathcal{E}}} \setminus \{\emptyset\}$. This frame represents agents with very low evidential demands. For these agents, having an argument for proposition P among their evidence is enough to justify P, regardless, e.g., whether the argument contradicts with the other available arguments.

2. The *strong denseness frame of justification* , denoted by $\boldsymbol{\mathcal{J}}^{SD}$, is the set of all arguments consistent with $\boldsymbol{\tau_{\mathcal{E}}}$, that is, the set of all dense open sets in S—see Definition 2.20. This frame represents agents with high evidential demands. They form degrees of beliefs only based on arguments which do not contradict with, i.e., cannot be refuted by, any other argument. Consequently, they form degrees of belief only in those propositions that do not contradict with any available argument. This notion of justification is proposed by Baltag et al. (2016, 2022).

Given a qualitative evidence frame $(S, \boldsymbol{\mathcal{E}})$ and a proposition P, we say that T is justification for P w.r.t. $\boldsymbol{\mathcal{J}}$ if and only if $T \subseteq P$ and $T \in \boldsymbol{\mathcal{J}}$. This perspective on justification substantially differs from the understanding of justification in traditional epistemology, where the main interest lies in defining the notion of justification for belief, or answering the question "what justifies belief?". Our use of the term justification is more pragmatically motivated and it is simply intended to discern an agent's any evidence-based argument from the ones they actually see fit to support their beliefs. The following example shows how different these two frames of justification can look. For an example in context, see Section 3.3.

Example 3.1 Frames of Justification

Let $S = \{a, b, c, d\}$ and $\boldsymbol{\mathcal{E}} = \Big\{\{a, b\}, \{b, c\}, \{c, d\}\Big\}$ be a set of possible states and qualitative evidence set respectively.

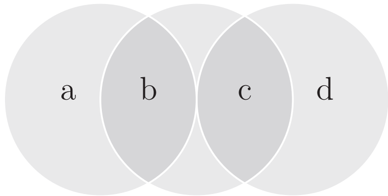


Figure 3.1: Representation of pieces of evidence by a Venn diagram.

According to the previous definitions, the Dempster-Shafer frame of justification for qualitative evidence frame $(S, \boldsymbol{\mathcal{E}})$ is

$$\boldsymbol{\mathcal{J}}^{DS} = \Big\{\{a, b\}, \{b, c\}, \{c, d\}, \{b\}, \{c\}, \{a, b, c\}, \{b, c, d\}, S\Big\}$$

while the strong denseness frame of justification is

$$\boldsymbol{\mathcal{J}}^{SD} = \Big\{\{b, c\}, \{a, b, c\}, \{b, c, d\}, S\Big\}.$$

Throughout this thesis, we will restrict our attention to Dempster-Shafer and strong denseness frames of justification. However, it is significant to note that these two frames are part of a larger family of frames of justification rooted in the notion of denseness. While $\boldsymbol{\mathcal{J}}^{DS}$ does not require an argument $A \in \boldsymbol{\tau_{\mathcal{E}}} \backslash \{\emptyset\}$ to be consistent with any other argument, $\boldsymbol{\mathcal{J}}^{SD}$ requires A to be consistent with all of them. We can define "intermediate" consistency requirements as well. Let us define the family of frames of justification

$$\boldsymbol{\mathcal{J}}^{\mathrm{Den}(k)} = \{A \in \boldsymbol{\tau_{\mathcal{E}}} \setminus \{\emptyset\} | A \cap \bigcap \boldsymbol{E} \neq \emptyset \text{ for all } \boldsymbol{E} \subseteq \boldsymbol{\mathcal{E}} \text{ such that } |\boldsymbol{E}| \leq k\},$$

where $|\boldsymbol{E}|$ is the number of elements in $\boldsymbol{E}$. For each k, $\boldsymbol{\mathcal{J}}^{\mathrm{Den}(k)}$ is formed by those $A \in \boldsymbol{\tau_{\mathcal{E}}} \backslash \{\emptyset\}$ to be consistent with the intersection of every c-combination of elements of $\boldsymbol{\mathcal{E}}$ with $c \leq k$. This way, as k increases, $\boldsymbol{\mathcal{J}}^{\mathrm{Den}(k)}$ converges to $\boldsymbol{\mathcal{J}}^{SD}$. We name $\boldsymbol{\mathcal{J}}^{\mathrm{Den}(k)}$ *depth-k denseness frame of justification*. For example, if we have $\boldsymbol{\mathcal{E}} = \{E_1, E_2, E_3\}$, the elements of $\boldsymbol{\mathcal{J}}^{\mathrm{Den}(1)}$ will have non-empty intersections with E_1, E_2 and E_3. The elements of $\boldsymbol{\mathcal{J}}^{\mathrm{Den}(2)}$ will will have non-empty intersections with E_1, E_2, E_3, $E_1 \cap E_2$, $E_1 \cap E_3$ and $E_2 \cap E_3$. The elements of $\boldsymbol{\mathcal{J}}^{\mathrm{Den}(3)}$ will will have non-empty intersections with E_1, E_2, E_3, $E_1 \cap E_2$, $E_1 \cap E_3$, $E_2 \cap E_3$ and

$E_1 \cap E_2 \cap E_3$. And $\boldsymbol{\mathcal{J}}^{\mathrm{Den}(3)}$ is actually $\boldsymbol{\mathcal{J}}^{SD}$ for $\boldsymbol{\mathcal{E}}$. Attending to this definition, agents could set their evidential demands on a scale of cautiousness related to the consistency of their arguments. Needless to say, other families rooted in different properties can be defined too. We leave the proper treatment of these other frames of justification for future work.

3.1.2 Quantitative Layer

The aim of this layer is to combine the degrees of certainty of the basic pieces of evidence via a mass function defined on the set of basic pieces of evidence $\boldsymbol{\mathcal{E}}$—see Definition 2.26. This way we obtain certainty values for every combination of pieces of evidence and these values sum up to one.

Let $(S, \boldsymbol{\mathcal{E}}^Q)$ be a quantitative evidence frame with ℓ pieces of evidence, i.e., $\boldsymbol{\mathcal{E}}^Q = \{(E_1, p_1), \ldots, (E_\ell, p_\ell)\}$. In Equation 3.1, we define a function $\delta : 2^{\boldsymbol{\mathcal{E}}} \to [0,1]$ to merge the certainty values p_j for $j = 1, \ldots, \ell$.

$$\delta(\boldsymbol{E}) = \prod_{E_j \in \boldsymbol{E}} p_j \prod_{E_j \notin \boldsymbol{E}} 1 - p_j \tag{3.1}$$

This function is a mass function defined over $\boldsymbol{\mathcal{E}}$ in accordance with Definition 2.26.

Proposition 3.1 *Given a quantitative evidence frame $(S, \boldsymbol{\mathcal{E}}^Q)$, the function δ defined in Equation (3.1) is a mass function over $\boldsymbol{\mathcal{E}}$.*

Proof.

Given $\boldsymbol{E} \in 2^{\boldsymbol{\mathcal{E}}}$, the function δ is well-defined and the value $\delta(\boldsymbol{E})$ is between 0 and 1 for $p_j \in (0,1)$ for all $j = 1, \ldots, \ell$. In addition, the total sum of $\delta(\boldsymbol{E})$ for all $\boldsymbol{E} \in 2^{\boldsymbol{\mathcal{E}}}$ is 1. Let $\boldsymbol{\mathcal{E}}_\lambda$ be the subset of $\boldsymbol{\mathcal{E}}$ formed by its first λ elements. Now, let us apply induction on the number of pieces of evidence in $\boldsymbol{\mathcal{E}}^Q$. For $\lambda = 1$, $2^{\boldsymbol{\mathcal{E}}_1} = \{\emptyset, \{E_1\}\}$ and

$$\sum_{\boldsymbol{E} \in 2^{\boldsymbol{\mathcal{E}}_1}} \delta(\boldsymbol{E}) = \prod_{E_j \in \emptyset} p_j \prod_{E_j \notin \emptyset} (1-p_j) \quad + \prod_{E_j \in \{E_1\}} p_j \prod_{E_j \notin \{E_1\}} (1-p_j) = 1-p_1+p_1 = 1.$$

Let us assume that $\sum_{\boldsymbol{E} \subseteq \boldsymbol{\mathcal{E}}_{\ell-1}} \delta(\boldsymbol{E}) = 1$. Given ℓ pieces of evidence in $\boldsymbol{\mathcal{E}}$, let us consider the partition $2^{\boldsymbol{\mathcal{E}}_\ell} = 2^{\boldsymbol{\mathcal{E}}_{\ell-1}} \cup \left\{\boldsymbol{E} \cup \{E_\ell\} | \boldsymbol{E} \in 2^{\boldsymbol{\mathcal{E}}_{\ell-1}}\right\}$ of $2^{\boldsymbol{\mathcal{E}}_\ell}$. Then, by the inductive hypothesis, the sum $\sum_{\boldsymbol{E} \in 2^{\boldsymbol{\mathcal{E}}_\ell}} \delta(\boldsymbol{E})$ is equal to

$$\sum_{\boldsymbol{E} \in 2^{\boldsymbol{\mathcal{E}}_{\ell-1}}} \Big((1-p_\ell) \prod_{E_j \in \boldsymbol{E}} p_j \prod_{E_j \notin \boldsymbol{E}} (1-p_j)\Big) + \sum_{\boldsymbol{E} \in 2^{\boldsymbol{\mathcal{E}}_{\ell-1}}} \Big(p_\ell \prod_{E_j \in \boldsymbol{E}} p_j \prod_{E_j \notin \boldsymbol{E}} (1-p_j)\Big),$$

which is equal to $(1 - p_\ell) \cdot 1 + p_\ell \cdot 1 = 1$. □

Intuitively, the function δ distributes the degree of certainty of a piece of evidence $E \in \boldsymbol{\mathcal{E}}$ among all the possible occurrences of E in presence of other pieces of evidence. To illustrate, consider a quantitative evidence frame $(S, \boldsymbol{\mathcal{E}}^Q)$ where $\boldsymbol{\mathcal{E}}^Q = \{(E_1, p_1), (E_2, p_2), (E_3, p_3)\}$. We then understand $\delta(\{E_1, E_2\}) = p_1 \cdot p_2 \cdot (1 - p_3)$ as the induced certainty of the proposition 'at least E_1 and E_2 are true' in the context of having exactly three pieces of evidence. This intuition is supported by the fact that, given a quantitative evidence frame $(S, \boldsymbol{\mathcal{E}}^Q)$ and δ defined as in equation (3.1), we have, for every $E_j \in \boldsymbol{\mathcal{E}}$, that the total sum of $\delta(\boldsymbol{E})$ for $\boldsymbol{E} \subseteq \boldsymbol{\mathcal{E}}$ such that $E_j \in \boldsymbol{E}$ is equal to p_j.

Proposition 3.2 *Given a quantitative evidence frame $(S, \boldsymbol{\mathcal{E}}^Q)$ and δ defined as in equation (3.1), we have, for every $E_j \in \boldsymbol{\mathcal{E}}$, that*

$$\sum_{\substack{\boldsymbol{E} \subseteq \boldsymbol{\mathcal{E}}: \\ E_j \in \boldsymbol{E}}} \delta(\boldsymbol{E}) = p_j.$$

Proof.

Without loss of generality, let us prove the result for $E_j = E_1$. First, let us notice that

$$\sum_{\substack{\boldsymbol{E} \subseteq \boldsymbol{\mathcal{E}}: \\ E_1 \in \boldsymbol{E}}} \sum \delta(\boldsymbol{E}) = \sum_{\substack{\boldsymbol{E} \subseteq \boldsymbol{\mathcal{E}}: \\ E_1 \in \boldsymbol{E}}} \Bigg(p_1 \prod_{\substack{E_k \in \boldsymbol{E}: \\ E_k \neq E_1}} p_k \prod_{\substack{E_k \notin \boldsymbol{E}: \\ E_k \neq E_1}} (1 - p_k) \Bigg)$$

In other words,

$$\sum_{\substack{\boldsymbol{E} \subseteq \boldsymbol{\mathcal{E}}: \\ E_1 \in \boldsymbol{E}}} \delta(\boldsymbol{E}) = p_1 \cdot \sum_{\boldsymbol{E} \subseteq \boldsymbol{\mathcal{E}} \setminus \{E_1\}} \Bigg(\prod_{E_k \in \boldsymbol{E}} p_k \prod_{\substack{E_k \notin \boldsymbol{E}, \\ E_k \neq E_1}} (1 - p_k) \Bigg) \tag{3.2}$$

Defining the set of evidence $\boldsymbol{\mathcal{E}}' = \boldsymbol{\mathcal{E}} \setminus \{E_1\}$, we have that

$$\sum_{\boldsymbol{E} \subseteq \boldsymbol{\mathcal{E}}'} \Bigg(\prod_{E_k \in \boldsymbol{E}} p_k \prod_{\substack{E_k \notin \boldsymbol{E}, \\ E_k \neq E_1}} (1 - p_k) \Bigg) = \sum_{\boldsymbol{E} \subseteq \boldsymbol{\mathcal{E}}'} \delta(\boldsymbol{E}),$$

where δ is defined over $\boldsymbol{\mathcal{E}}'$ now. By Proposition 3.1, $\sum_{\boldsymbol{E} \subseteq \boldsymbol{\mathcal{E}}'} \delta(\boldsymbol{E}) = 1$ and Equation (3.2) is equal to p_1. □

Nevertheless, it is easier to read δ as a *system* of certainty assignments rather than trying to interpret each $\delta(\boldsymbol{E})$ epistemically, since it assigns values to every combination of pieces of evidence. It is merely a method to distribute the certainty of a single piece of evidence to all the ways it can be observed in combination with the other pieces of evidence. We will illustrate how this function works in the following example.

Example 3.2 Function δ

Let $S = \{a, b, c, d\}$ and $\boldsymbol{\mathcal{E}}^Q = \left\{E_1, E_2, E_3\right\}$ be a set of possible states and quantitative evidence set, respectively, where $E_1 = (\{a, b\}, 0.6)$, $E_2 = (\{b, c\}, 0.8)$, and $E_3 = (\{c, d\}, 0.7)$. The image of function δ applied on the quantitative evidence frame $(S, \boldsymbol{\mathcal{E}}^Q)$ is:

$$\begin{array}{lcllcl}
\delta(\emptyset) & = & 0.02 & \delta(\{E_1, E_2\}) & = & 0.14 \\
\delta(\{E_1\}) & = & 0.04 & \delta(\{E_1, E_3\}) & = & 0.08 \\
\delta(\{E_2\}) & = & 0.10 & \delta(\{E_2, E_3\}) & = & 0.22 \\
\delta(\{E_3\}) & = & 0.06 & \delta(\{E_1, E_2, E_3\}) & = & 0.34
\end{array}$$

Table 3.1: Image of the function δ for $\boldsymbol{\mathcal{E}}^Q = \{E_1, E_2, E_3\}$.

If we were to consider $\boldsymbol{\mathcal{E}}^Q = \left\{E_1, E_2\right\}$ with $E_1 = (\{a, b\}, 0.6)$, $E_2 = (\{b, c\}, 0.8)$ instead, the image of function δ would be:

$$\begin{array}{lcl}
\delta(\{\emptyset\}) & = & 0.08 \\
\delta(\{E_1\}) & = & 0.12 \\
\delta(\{E_2\}) & = & 0.32 \\
\delta(\{E_1, E_2\}) & = & 0.48
\end{array}$$

Table 3.2: Image of the function δ for $\boldsymbol{\mathcal{E}}^Q = \{E_1, E_2\}$.

So representing the certainty of evidence by the δ function enables us to capture the presence or absence of pieces of evidence before turning these certainty values into belief—i.e., before taking into account other kinds of information about the agent, such as the frame of justification, for example.

3.1.3 Bridging Layer

Having introduced the qualitative and quantitative layers, we can now connect these two to calculate degrees of beliefs based on a quantitative evidence frame $(S, \boldsymbol{\mathcal{E}}^Q)$ and a given frame of justification $\boldsymbol{\mathcal{J}}$. To this end, we start the section by defining a family of functions to map $2^{\boldsymbol{\mathcal{E}}}$—the domain of the mass function δ defined in equation (3.1)—to $\boldsymbol{\tau}_{\boldsymbol{\mathcal{E}}}$. Finally, we will define two mass functions, one with domain $\boldsymbol{\tau_{\mathcal{E}}}$ and another one with domain $\boldsymbol{\mathcal{J}}$, that will be used to compute the degrees of belief our final model returns.

Evidence Allocation Functions

At this point, we have a set of justifications $\boldsymbol{\mathcal{J}}$ and a merged measure of certainty distributed over the elements of $2^{\boldsymbol{\mathcal{E}}}$. Our goal is to link the mass values defined for the elements of $2^{\boldsymbol{\mathcal{E}}}$ by δ to the elements of $\boldsymbol{\tau_{\mathcal{E}}}$ and, in turn, to the elements of $\boldsymbol{\mathcal{J}}$. Mapping these two sets is not a trivial issue as there are many ways to do so. For example, given two pieces of evidence E_1 and E_2 in $\boldsymbol{\mathcal{E}}$, the δ-value associated with the set $\{E_1, E_2\}$ could be mapped to different elements of $\boldsymbol{\tau_{\mathcal{E}}}$ depending on the interpretation the agent gives to these values. A strict interpretation could state that the agent assumes that every piece of evidence in the set $\{E_1, E_2\}$ contains the actual world, which would map $\{E_1, E_2\}$ to $E_1 \cap E_2$. Conversely, a moderate interpretation could state that the agent assumes that at least one of the elements of $\{E_1, E_2\}$ contains the actual world, which would map $\{E_1, E_2\}$ to $E_1 \cup E_2$. These interpretations depend on the context of the agent, as it is the degree of reliability of the sources that will make the agent establish some assumptions or others. We capture various ways of interpreting the mass values provided by δ via *evidence allocation functions*.

Definition 3.1 (Set of evidence allocation functions). Let $(S, \boldsymbol{\mathcal{E}})$ be a qualitative evidence frame. A *set of evidence allocation functions* $\mathfrak{F}$ on $(S, \boldsymbol{\mathcal{E}})$ is a set of functions from $2^{\boldsymbol{\mathcal{E}}}$ to $\boldsymbol{\tau_{\mathcal{E}}}$ (the topology generated by $\boldsymbol{\mathcal{E}}$) such that for all $f, g \in \mathfrak{F}$:

1. $f(\emptyset) = S$,
2. for all non-empty $\boldsymbol{E} \subseteq \boldsymbol{\mathcal{E}}$, $f(\boldsymbol{E}) \in \boldsymbol{\tau_E}$—the topology generated by $\boldsymbol{E}$—and it is dense in $\bigcup \boldsymbol{E}$ w.r.t. $\boldsymbol{\tau_E}$; or $f(\boldsymbol{E}) = \emptyset$, and
3. for all $\boldsymbol{E} \subseteq \boldsymbol{\mathcal{E}}$ and every f, g in $\mathfrak{F}$, $f(\boldsymbol{E}) \subseteq g(\boldsymbol{E})$ or $g(\boldsymbol{E}) \subseteq f(\boldsymbol{E})$.

Note that since $\boldsymbol{E} \subseteq \boldsymbol{\mathcal{E}}$, we have $\boldsymbol{\tau_E} \subseteq \boldsymbol{\tau_{\mathcal{E}}}$—it follows by Definition 2.15.

In what follows, we assume that all evidence allocation functions are defined on a qualitative evidence frame $(S, \boldsymbol{\mathcal{E}})$ with ℓ pieces of evidence and omit mention of it. The first item of this definition preserves the notion of uncertainty since in our context it is modeled by associating the value $1 - p_j$ (for $j = 1, \ldots, \ell$) with the total set. On the other hand, conditions 2 and 3 establish some minimal rationality constraints. Condition 2 states that an evidence allocation function f assigns $\boldsymbol{E}$ to an argument that is generated by $\boldsymbol{E}$ and that does not contradict any other argument generated by $\boldsymbol{E}$. So, an evidence allocation function does not assign a set of evidence $\boldsymbol{E}$ to some argument that cannot be produced within $\boldsymbol{E}$ or that is inconsistent with $\boldsymbol{E}$. Condition 3 ensures that two agents with the same set of evidence allocation functions will associate $\boldsymbol{E}$ with arguments such that one entails the other. In this sense, $\boldsymbol{E}$ cannot pull these agents in different directions with respect to their evidence. Let us see three examples of these functions.

Proposition 3.3 *Given a set of evidence allocation functions* $\mathfrak{F}$ *and the function* $i : 2^{\boldsymbol{\mathcal{E}}} \to \boldsymbol{\tau_{\mathcal{E}}}$ *such that*

$$i(\boldsymbol{E}) = \begin{cases} \bigcap \boldsymbol{E} & \text{if } \boldsymbol{E} \neq \emptyset, \\ S & \text{otherwise} \end{cases}$$

the set $\mathfrak{F} \cup \{i\}$ *is a set of evidence allocation functions.*

Proof.

> First, every element of $\mathfrak{F}\cup\{i\}$ maps the empty set to S by definition. Secondly, recall that the topology $\boldsymbol{\tau_E}$ is generated by the element of $\boldsymbol{E}$ by closing it under finite intersections and arbitrary unions. This implies that $i(\boldsymbol{E}) \in \boldsymbol{\tau_E}$ and $i(\boldsymbol{E}) = \bigcap \boldsymbol{E} \subseteq T$ for all $T \in \boldsymbol{\tau_E} \setminus \{\emptyset\}$—i.e., $i(\boldsymbol{E})$ is the smallest element of $\boldsymbol{\tau_E}$. Consequently, $i(\boldsymbol{E})$ is either the empty set or it is a dense element of $\boldsymbol{\tau_E}$. In addition, given $f \in \mathfrak{F}$, $f(\boldsymbol{E}) \in \boldsymbol{\tau_E}$—by Definition 3.1. So, if $f(\boldsymbol{E}) \neq \emptyset$, then $i(\boldsymbol{E}) \subseteq f(\boldsymbol{E})$. □

Proposition 3.4 *Given a set of evidence allocation functions* $\mathfrak{F}$ *and the function* $u : 2^{\boldsymbol{\mathcal{E}}} \to \boldsymbol{\tau_{\mathcal{E}}}$ *such that*

$$u(\boldsymbol{E}) = \begin{cases} \bigcup \boldsymbol{E} & \text{if } \boldsymbol{E} \neq \emptyset, \\ S & \text{otherwise} \end{cases}$$

the set $\mathfrak{F} \cup \{u\}$ *is a set of evidence allocation functions.*

Proof.

> Conditions 1 and 3 of Definition 3.1 follows from similar arguments as in the proof of Proposition 3.3. In addition, considering $\boldsymbol{E} \in 2^{\boldsymbol{\mathcal{E}}}$, the topology $\boldsymbol{\tau_E}$ is a set generated by the elements of $\boldsymbol{E}$ by closing it under finite intersections and arbitrary unions. This implies that $u(\boldsymbol{E}) = \bigcup \boldsymbol{E}$ is the largest element of $\boldsymbol{\tau_E}$. Therefore, for all $T \in \boldsymbol{\tau_E} \setminus \{\emptyset\}$, $T \cap u(\boldsymbol{E}) = T \neq \emptyset$, meaning that $u(\boldsymbol{E})$ is also dense in $\bigcup \boldsymbol{E}$ w.r.t. $\boldsymbol{\tau_E}$. □

If there exists a function f such that $u(\boldsymbol{E}) \subset f(\boldsymbol{E})$ or $f(\boldsymbol{E}) \subset i(\boldsymbol{E})$, then f violates Condition 2 of Definition 3.1. Therefore, we infer the following corollary, showing the boundaries of possible evidence allocation functions.

Corollary 3.1 *Given a set of evidence allocation functions* $\mathfrak{F}$, *a function* $f \in \mathfrak{F}$, *and* $\boldsymbol{E} \in 2^{\boldsymbol{\mathcal{E}}}$, *we have* $i(\boldsymbol{E}) \subseteq f(\boldsymbol{E}) \subseteq u(\boldsymbol{E})$.

For our last example, we need the following auxiliary lemma.

Lemma 3.1 *Let $(S, \boldsymbol{\mathcal{E}})$ be a qualitative evidence frame, $\boldsymbol{E} \in 2^{\boldsymbol{\mathcal{E}}}$, and dense$(\boldsymbol{E})$ be the set of dense elements of $\boldsymbol{\tau_E}$ in $\bigcup \boldsymbol{E}$. Then the order (dense$(\boldsymbol{E}), \subseteq)$ has a minimum, i.e., there is an element $M \in$ dense$(\boldsymbol{E})$ such that $M \subseteq T$ for every $T \in$ dense$(\boldsymbol{E})$.*

Proof.

Let $\mathbb{M}$ be the set of all $\boldsymbol{A} \in 2^{\boldsymbol{E}}$ such that (1) $\bigcap \boldsymbol{A} \neq \emptyset$, and (2) if $\boldsymbol{A}' \in 2^{\boldsymbol{E}}$ and $\boldsymbol{A} \subset \boldsymbol{A}'$, then $\bigcap \boldsymbol{A}' = \emptyset$. We will say that a set $\boldsymbol{A} \in 2^{\boldsymbol{E}}$ which satisfies (1) and (2) is a maximal set with non-empty intersection. Then

$$\min\Big((\text{dense}(\boldsymbol{E}), \subseteq)\Big) = \bigcup\{\bigcap \boldsymbol{A} | \boldsymbol{A} \in \mathbb{M}\}.$$

To simplify notation, let $M := \bigcup\{\bigcap \boldsymbol{A} | \boldsymbol{A} \in \mathbb{M}\}$. First, M is in $\boldsymbol{\tau_E}$ by the definition of generated topology since every $\boldsymbol{A} \in \mathbb{M}$ only contains elements of $\boldsymbol{E}$. In addition, M is dense w.r.t. $\boldsymbol{\tau_E}$. Any element $T \in \boldsymbol{\tau_E}$ forms part of a maximal set with non-empty intersection: if T has empty intersections with all other elements of $\boldsymbol{\tau_E}$, then $\{T\}$ is a maximal set with non-empty intersection. Since M contains the intersections of all these maximal sets, in particular, $M \cap T \neq \emptyset$.

Now, let us prove that $M \subseteq T$ for every element $T \in$ dense$(\boldsymbol{E})$. Let $\boldsymbol{A}$ be an arbitrary element of $\mathbb{M}$ and T an arbitrary element of dense$(\boldsymbol{E})$. Since $\boldsymbol{A}$ is a subset of $\boldsymbol{E}$, $\bigcap \boldsymbol{A} \in \boldsymbol{\tau_E}$. In addition, T being dense in $\boldsymbol{\tau_E}$ implies that $T \cap \bigcap \boldsymbol{A} \neq \emptyset$. Let us take the subset of $\boldsymbol{E}$ formed by every element of $\boldsymbol{E}$ which contains $T \cap \bigcap \boldsymbol{A}$. This set is non-empty since it includes at least $\boldsymbol{A}$. If this set contains a piece of evidence $E \in \boldsymbol{E}$ such that $E \notin \boldsymbol{A}$, then $\bigcap \boldsymbol{A} \cap \{E\}$ is non-empty and $\boldsymbol{A}$ does not satisfy the maximal finite intersection property. Therefore, every piece of evidence E in $\boldsymbol{E}$ that contains $T \cap \bigcap \boldsymbol{A}$ is in $\boldsymbol{A}$. By definition of topology generated by $\boldsymbol{E}$, if every element in $\boldsymbol{E}$ that contains a also contains b, then every element of $\boldsymbol{\tau_E}$ that contains a contains b too. We have shown that every element of $\boldsymbol{E}$ that contains $T \cap \bigcap \boldsymbol{A}$ also contains $\bigcap \boldsymbol{A}$. So every element of $\boldsymbol{\tau}_E$ that contains $T \cap \bigcap \boldsymbol{A}$, for example T, also contains $\bigcap \boldsymbol{A}$. Hence $M \subseteq T$ for any $T \in$ dense$(\boldsymbol{E})$, that is, M is the minimum of the dense sets in $\boldsymbol{\tau_E}$ with the subset relation. □

Proposition 3.5 *Given a set of evidence allocation functions $\mathfrak{F}$ and the function $d : 2^{\boldsymbol{\mathcal{E}}} \to \boldsymbol{\tau_{\mathcal{E}}}$ such that*

$$d(\boldsymbol{E}) = \begin{cases} min\Big((dense(\boldsymbol{E}), \subseteq)\Big) & \text{if } \boldsymbol{E} \neq \emptyset, \\ S & \text{otherwise} \end{cases}$$

the set $\mathfrak{F} \cup \{d\}$ is a set of evidence allocation functions.

Proof.

> The first two conditions of Definition 3.1 hold by definition. In addition, for any $f \in \mathfrak{F}$, we know $f(\boldsymbol{E}) = \emptyset$ or $f(\boldsymbol{E})$ is dense w.r.t. $\boldsymbol{\tau}_{\boldsymbol{E}}$. In the former case, $f(\boldsymbol{E}) \subseteq d(\boldsymbol{E})$. In the latter one, $d(\boldsymbol{E}) \subseteq f(\boldsymbol{E})$ by Lemma 3.1. □

Corollary 3.2 *The set of functions $\mathfrak{F} = \{i, u, d\}$ is a set of evidence allocation functions.*

As we mentioned earlier, evidence allocation functions can be understood as contextual parameters. Depending on the context, agents may prefer to place the available information in weaker or stronger arguments—arguments which contain more or fewer elements respectively—to avoid increasing the uncertainty of the model—as it can happen by mapping elements of $2^{\boldsymbol{\mathcal{E}}}$ to the total set S—, or to avoid discarding information—as it can happen by mapping elements of $2^{\boldsymbol{\mathcal{E}}}$ to the empty set. In this regard, we have already seen that an agent who assumes that every piece of evidence in the set $\{E_1, E_2\}$ contains the actual world may use function i; and that an agent who assumes that at least one piece of evidence in the set $\{E_1, E_2\}$ contains the actual world may use function u. Similarly, function d would be a good option for those agents that assume that the actual world is consistent with the strongest argument that E_1 and E_2 can produce together. That is, if we have $E_1 = \{a, b\}$, $E_2 = \{b, c\}$, the strongest argument that E_1 and E_2 can produce together is $\{b\}$. If we add $E_3 = \{c, d\}$, the strongest argument that E_1, E_2 and E_3 can produce together is $\{b, c\}$. In terms of reliability, this last case can be described as *the combination of all sources is reliable*. To foster intuitions about this combined reliability, we can think about a coloured object measured by three sensors: a red light sensor, a green light sensor and a blue light sensor. Even if their measurements are correct, drawing conclusions based only on one or two of the sensors will be incorrect in most cases.

Belief Functions

Now that we know how to link the power set of $\boldsymbol{\mathcal{E}}$ and $\boldsymbol{\tau}_{\boldsymbol{\mathcal{E}}}$, let us define some functions that will allow us to compute degrees of belief given a body of possibly mutually inconsistent, partial and uncertain evidence as well as an agent's evidential demands.

Let $(S, \boldsymbol{\mathcal{E}}^Q)$ be a quantitative evidence frame and $\mathfrak{F}$ a set of evidence allocation functions. Then, we define the function $\delta_{\boldsymbol{\tau}} : \mathfrak{F} \times 2^S \to [0, 1]$ as:

$$\delta_{\boldsymbol{\tau}}(f, T) = \begin{cases} \sum\limits_{\boldsymbol{E}: f(\boldsymbol{E}) = T} \delta(\boldsymbol{E}) & \text{if } T \in \boldsymbol{\tau}_{\boldsymbol{\mathcal{E}}}, \\ 0 & \text{otherwise.} \end{cases} \tag{3.3}$$

Fixing $f \in \mathfrak{F}$, $\delta_{\mathcal{T}}(f, \cdot)$ is a mass function over S since δ is a mass function over $\boldsymbol{\mathcal{E}}$.

By normalizing the previous function with respect to a frame of justification $\boldsymbol{\mathcal{J}}$, we define the function $\delta_{\mathcal{J}} : \mathfrak{F} \times 2^S \to [0, 1]$ as:

$$\delta_{\mathcal{J}}(f, A) = \begin{cases} \dfrac{\delta_{\mathcal{T}}(f, A)}{\sum_{T \in \mathcal{J}} \delta_{\mathcal{T}}(f, T)} & \text{if } A \in \boldsymbol{\mathcal{J}}, \\ 0 & \text{otherwise.} \end{cases} \tag{3.4}$$

The following result ensures that for every frame of justification $\boldsymbol{\mathcal{J}}$ and every evidence allocation function f, the function $\delta_{\mathcal{J}}(f, \cdot)$ can be used to obtain degrees of belief additively.

Proposition 3.6 *Given a quantitative evidence frame* $(S, \boldsymbol{\mathcal{E}}^Q)$*, a frame of justification* $\boldsymbol{\mathcal{J}}$*, and an evidence allocation function* $f \in \mathfrak{F}$*, the function* $\delta_{\mathcal{J}}(f, \cdot)$ *defined in Equation* (3.4) *is a mass function over* S *such that* $\delta_{\mathcal{J}}(f, \emptyset) = 0$.

Proof.

> The above fraction is well-defined for every non-dogmatic evidence set—that is, for $\boldsymbol{\mathcal{E}}^Q = \{(E_j, p_j)\}_{j=1,\dots,\ell}$ such that $p_j \neq 1$ for all j—since the set of possible states S belongs to every frame of justification. In addition, $\delta_{\mathcal{J}}(f, \emptyset) = 0$ since $\emptyset \notin \boldsymbol{\mathcal{J}}$ for the definition of argument. Finally, the total sum of its values is equal to 1 because $\delta_{\mathcal{T}}(f, \cdot)$ is a mass function over S. □

At last, we can define the *degree of belief* for proposition $P \subseteq S$ via a multi-layer belief function given a quantitative evidence frame, a frame of justification, and a set of evidence allocation functions.

Definition 3.2 (Multi-layer belief function). Let $(S, \boldsymbol{\mathcal{E}}^Q)$ be a quantitative evidence frame, $\boldsymbol{\mathcal{J}}$ a frame of justification and $\mathfrak{F}$ a set of evidence allocation functions. We call a function $\mathrm{Bel}_{\mathcal{J}} : \mathfrak{F} \times 2^S \to [0, 1]$ defined as

$$\mathrm{Bel}_{\mathcal{J}}(f, P) = \sum_{A \subseteq P} \delta_{\mathcal{J}}(f, A)$$

a *multi-layer belief function.* When $\boldsymbol{\mathcal{J}}$ and f are clear from context, we omit mention of them and write $\mathrm{Bel}(P)$.

Proposition 3.7 *Given a quantitative evidence frame* $(S, \boldsymbol{\mathcal{E}}^Q)$*, a frame of justification* $\boldsymbol{\mathcal{J}}$*, and an evidence allocation function* f*, the function* $\mathrm{Bel}_{\mathcal{J}}(f, \cdot)$ *defined in Definition 3.2 is a belief function according to Definition 2.5.*

Proof.

By Proposition 2.5 and Proposition 3.6, the function $\mathrm{Bel}_{\mathcal{J}}(f, \cdot)$ is a belief function (Shafer, 1976, p. 51). $\square$

3.2 Relation to Other Belief Models

The multi-layer belief model aims to extend Dempster's rule of combination and topological models of evidence by adding features that their original settings do not have. Specifically, it expands Dempster's rule of combination by distinguishing between evidence and justifications, and topological models of evidence by representing degrees of certainty. In this section, we will demonstrate that this model properly extends them, since both Dempster's rule of combination and topological models of evidence can be represented as multi-layer belief models. Furthermore, we show that the multi-layer belief model also captures the transferable belief model to some extent, and offers an alternative both to the transferable belief model and the disjunctive rule of combination.

We start proving that, given a quantitative evidence frame $(S, \mathcal{E}^Q)$ with ℓ pieces of evidence, running the multi-layer belief model with Dempster-Shafer frame of justification and evidence allocation function i, we obtain exactly the belief function given by Dempster's rule of combination.

Proposition 3.8 *Let $(S, \mathcal{E}^Q)$ be a quantitative evidence frame, ℓ the number of elements in $\mathcal{E}$, $i : 2^{\mathcal{E}} \to \tau_{\mathcal{E}}$ the evidence allocation function defined in Proposition 3.3, and $\mathcal{J}^{DS}$ Dempster-Shafer frame of justification, that is, $\mathcal{J}^{DS} = \tau_{\mathcal{E}} \setminus \{\emptyset\}$. Let us consider the belief function $\mathrm{Bel}_{\mathcal{J}^{DS}}(i, \cdot) : 2^S \to [0,1]$ defined by the multi-layer belief model and bel $: 2^S \to [0,1]$ a belief function obtained by applying Dempster's rule of combination to the basic belief assignments $\{m_j | j = 1, \ldots, \ell\}$ such that $m_j(E_j) = p_j$, $m_j(S) = 1 - p_j$ for every $j = 1, \ldots, \ell$. Then,*

$$\mathrm{Bel}_{\mathcal{J}^{DS}}(i, P) = bel(P)$$

for every $P \subseteq S$.

Proof.

Let us consider $m = \oplus_j m_j$. Given a proposition $P \subseteq S$, $\mathrm{Bel}_{\mathcal{J}^{DS}}(i, P) = \sum_{A \subseteq P} \delta_{\mathcal{J}^{DS}}(i, A)$ and $\mathrm{bel}(P) = \sum_{A \subseteq P} m(A)$. So let us prove that $\delta_{\mathcal{J}^{DS}}(i, A) = m(A)$ to prove the claim.

On the one hand, for $A \in \mathcal{J}^{DS}$

$$\delta_{\mathcal{J}^{DS}}(i, A) = \frac{\delta_{\tau}(i, A)}{\sum_{T \in \mathcal{J}^{DS}} \delta_{\tau}(i, T)} \tag{3.5}$$

where $\boldsymbol{\mathcal{J}}^{DS} = \boldsymbol{\tau}_{\boldsymbol{\mathcal{E}}} \setminus \{\emptyset\}$ and $\delta_{\boldsymbol{\tau}}(i, A) = \sum_{\bigcap \boldsymbol{E}=A} \delta(\boldsymbol{E})$ with $\boldsymbol{E} \in 2^{\boldsymbol{\mathcal{E}}}$. On the other hand,

$$m(A) = \frac{\sum_{\bigcap_j A_j = A} m_1(A_1) \cdot \ldots \cdot m_\ell(A_\ell)}{\sum_{\bigcap_j B_j \neq \emptyset} m_1(B_1) \cdot \ldots \cdot m_\ell(B_\ell)}. \tag{3.6}$$

Let us see that $\sum_{\bigcap_j A_j = A} m_1(A_1) \cdot \ldots \cdot m_\ell(A_\ell) = \sum_{\bigcap \boldsymbol{E}=A} \delta(\boldsymbol{E})$.

The functions m_j are basic belief assignments such that $m_j(A) \neq 0$ only when $A = E_j$ or $A = S$. Therefore,

$$\sum_{\bigcap_j A_j = A} m_1(A_1) \cdot \ldots \cdot m_\ell(A_\ell) \quad = \sum_{\substack{A_j \in \{E_j, S\}: \\ \bigcap_j A_j = A}} m_1(A_1) \cdot \ldots \cdot m_\ell(A_\ell).$$

Let us define a family of vectors $\bar{b}$ of length k such that $b_j \in \{E_j, S\}$. Then, the previous formula is equal to

$$\sum_{\bar{b}: \bigcap_j b_j = A} m_1(b_1) \cdot \ldots \cdot m_\ell(b_\ell).$$

There is a one-to-one correspondence between the family of vectors $\bar{b}$ and $2^{\boldsymbol{\mathcal{E}}}$. For example, we can define the function $f : \{E_j, S\}^\ell \to 2^{\boldsymbol{\mathcal{E}}}$ such that $f(\bar{b}) = \{E_j | b_j \neq S\}$. This function is bijective and $\bigcap b_j = A$ if and only if $\bigcap f(\bar{b}) = A$. This means that for every vector $\bar{b}$ such that $\bigcap b_j = A$ there exists exactly one $\boldsymbol{E} \in 2^{\boldsymbol{\mathcal{E}}}$ such that $\bigcap \boldsymbol{E} = A$. Consequently, the numerator of equation (3.6) is equal to

$$\sum_{\bigcap \boldsymbol{E}=A} \Big(\prod_{E_j \in \boldsymbol{E}} m_j(E_j) \prod_{E_j \notin \boldsymbol{E}} m_j(S) \Big) = \sum_{\bigcap \boldsymbol{E}=A} \Big(\prod_{E_j \in \boldsymbol{E}} p_j \prod_{E_j \notin \boldsymbol{E}} (1-p_j) \Big) = \sum_{\bigcap \boldsymbol{E}=A} \delta(\boldsymbol{E}).$$

We also know that every B_j in the denominator of Equation (3.6) belongs to $\{E_j, S\}$, so we can consider the same family of vectors $\bar{b}$ and the one-to-one correspondence between this family and $2^{\boldsymbol{\mathcal{E}}}$. If we unfold the definition of the denominator of equation (3.5), we get:

$$\sum_{T \in \boldsymbol{\mathcal{J}}^{DS}} \delta_{\boldsymbol{\tau}}(i, T) = \sum_{T \in \boldsymbol{\tau}_{\boldsymbol{\mathcal{E}}} \setminus \{\emptyset\}} \Big(\sum_{\bigcap \boldsymbol{E}=T} \delta(\boldsymbol{E}) \Big) = \sum_{\bigcap \boldsymbol{E} \neq \emptyset} \Big(\prod_{E_j \in \boldsymbol{E}} p_j \prod_{E_j \notin \boldsymbol{E}} (1 - p_j) \Big)$$

By applying the bijective function f that was defined above, this formula is equal to

$$\sum_{\bar{b}:\bigcap b_j \neq \emptyset} m_1(b_1)\cdot\ldots\cdot m_\ell(b_\ell) \;=\; \sum_{\bigcap_j B_j \neq \emptyset} m_1(B_1)\cdot\ldots\cdot m_\ell(B_\ell),$$

which proves our statement. □

The multi-layer belief model can also return the same outcome as topological models of evidence. In this case, we must consider the strong denseness frame of justification and the minimum dense set allocation function d given in Proposition 3.5. Notice that topological models of evidence is a qualitative approach, so the only properties we will use of the p_j values are that they are positive and smaller than 1.

Proposition 3.9 *Given a quantitative evidence frame $(S, \boldsymbol{\mathcal{E}}^Q)$, ℓ the number of elements in $\boldsymbol{\mathcal{E}}$, and the evidence allocation function $d : 2^{\boldsymbol{\mathcal{E}}} \to \boldsymbol{\tau_{\mathcal{E}}}$ defined in Proposition 3.5, let us consider $\mathrm{Bel}_{\boldsymbol{\mathcal{J}}^{SD}}(d,\cdot) : 2^S \to [0,1]$ the belief function defined by the multi-layer belief model; and $B : 2^S \to \{0,1\}$ a belief operator such that $B(P) = 1$ if and only if there exists $D \in \boldsymbol{\tau_{\mathcal{E}}}$ such that $D \subseteq P$ and $D \cap T \neq \emptyset$ for all $T \in \boldsymbol{\tau_{\mathcal{E}}} \setminus \{\emptyset\}$. Then,*

$$B(P) = 1 \textit{ if and only if } \mathrm{Bel}_{\boldsymbol{\mathcal{J}}^{SD}}(d, P) > 0$$

for every $P \subseteq S$.

Proof.

Let us prove the left to right direction. By definition, $B(P) = 1$ if and only if there exists $D \in \boldsymbol{\tau_{\mathcal{E}}}$ such that $D \subseteq P$ and $D \cap T \neq \emptyset$ for all $T \in \boldsymbol{\tau_{\mathcal{E}}} \setminus \{\emptyset\}$. That is if $B(P) = 1$ then there is a set D contained in P which is in $\boldsymbol{\tau_{\mathcal{E}}}$ and is dense in it. Therefore, $d(\boldsymbol{\mathcal{E}}) \subseteq D$ since the topology generated by $\boldsymbol{\mathcal{E}}$ is exactly $\boldsymbol{\tau_{\mathcal{E}}}$ and the proof of Lemma 3.1 shows that $d(\boldsymbol{\mathcal{E}})$ is included in every element of dense$(\boldsymbol{\mathcal{E}})$. Consequently,

$$\mathrm{Bel}_{\boldsymbol{\mathcal{J}}^{SD}}(d, P) = \sum_{A \subseteq P} \delta_{\boldsymbol{\mathcal{J}}^{SD}}(d, A) \geq \delta_{\boldsymbol{\mathcal{J}}^{SD}}(d, d(\boldsymbol{\mathcal{E}}))$$

And $\delta_{\boldsymbol{\mathcal{J}}^{SD}}(d, d(\boldsymbol{\mathcal{E}})) > 0$ if and only if $d(\boldsymbol{\mathcal{E}}) \in \boldsymbol{\mathcal{J}}^{SD}$ and $\delta_{\boldsymbol{\tau}}(d, d(\boldsymbol{\mathcal{E}})) > 0$. Both conditions hold by definition: $\delta_{\boldsymbol{\tau}}(d, d(\boldsymbol{\mathcal{E}})) = \sum_{d(\boldsymbol{E})=d(\boldsymbol{\mathcal{E}})} \delta(\boldsymbol{E})$ which is greater or equal to $\delta(\boldsymbol{\mathcal{E}}) = p_1 \cdot \ldots \cdot p_\ell > 0$.

Now, let us see the right-left implication. If $\mathrm{Bel}_{\boldsymbol{\mathcal{J}}^{SD}}(d, P) > 0$ then there is a set $\boldsymbol{E} \in 2^{\boldsymbol{\mathcal{E}}}$ such that $d(\boldsymbol{E}) \subseteq P$ and $d(\boldsymbol{E}) \in \boldsymbol{\mathcal{J}}^{SD}$. This means that $d(\boldsymbol{E})$ is a dense element of $\boldsymbol{\tau_{\mathcal{E}}}$, so there exists $D \in \boldsymbol{\tau_{\mathcal{E}}}$ such that $D \subseteq P$ and $D \cap T \neq \emptyset$ for all $T \in \boldsymbol{\tau_{\mathcal{E}}} \setminus \{\emptyset\}$. This proves that $B(P) = 1$. □

These two previous results show that the multi-layer belief model is able to return different belief functions based on the same body of evidence. Concretely, every combination of the (non-exhaustive) list of frames of justification and evidence allocation functions presented in Section 3.1 will return a different belief function. We call this property *adaptability to the context*. We assert that the multi-layer belief model is adaptable to the context because it allows the agent to adjust its parameters according to the current situation, without having to change the model or the evidence. Let us explore other two examples.

A common alternative to Dempster's rule of combination in the context of (partially) non-reliable data sources is the disjunctive rule of combination—see Definition 2.12. This rule replicates Dempster's rule of combination but replaces the intersection by the union, and allows the empty set to have non-negative values. Therefore, it does not require normalization. A natural alternative to the disjunctive rule of combination within the multi-layer belief model is to fix the Dempster-Shafer frame of justification and the evidence allocation function u—see Proposition 3.4. In the following example we will see that the disjunctive rule and the multi-layer belief model for Dempster-Shafer frame of justification and evidence allocation function u, $\mathrm{Bel}_{\mathcal{J}^{DS}}(u,\cdot)$, are not equivalent, but both adhere to the scenario where *at least one of the sources is reliable*.

Example 3.3 Alternative to Disjunctive Rule of Combination

Let $S = \{a, b, c, d\}$ be a set of possible states and $\mathcal{E}^Q = \{E_1, E_2, E_3\}$ where $E_1 = (\{a, b\}, 0.6)$, $E_2 = (\{b, c\}, 0.8)$ and $E_3 = (\{c, d\}, 0.7)$.

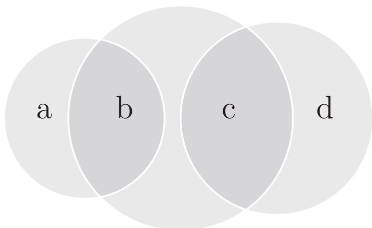


Figure 3.2: Representation of uncertain pieces of evidence by a Venn diagram. The size of the area represents the degree of certainty of the corresponding piece of evidence.

Fixing $\mathcal{J}^{DS}$ as frame of justification and u as evidence allocation function, we obtain:

$$\begin{aligned}
\delta_{\mathcal{J}^{DS}}(u, \{a, b\}) &= 0.036 & \delta_{\mathcal{J}^{DS}}(u, \{a, b, c\}) &= 0.144\\
\delta_{\mathcal{J}^{DS}}(u, \{b, c\}) &= 0.096 & \delta_{\mathcal{J}^{DS}}(u, \{b, c, d\}) &= 0.224\\
\delta_{\mathcal{J}^{DS}}(u, \{c, d\}) &= 0.056 & \delta_{\mathcal{J}^{DS}}(u, S) &= 0.444
\end{aligned}$$

If we apply the disjunctive rule of combination to this example, the output would collapse to S since $\{a, b\} \cup \{b, c\} \cup \{c, d\} = S$. Since this rule allows to associate non-zero values to the empty set, we can consider the following case instead:

$$
\begin{array}{rclrclrcl}
m_1(\emptyset) &=& 0.05, & m_1(\{a,b\}) &=& 0.6, & m_1(S) &=& 0.35\\
m_2(\emptyset) &=& 0.05, & m_2(\{b,c\}) &=& 0.8, & m_2(S) &=& 0.15\\
m_3(\emptyset) &=& 0.05, & m_3(\{c,d\}) &=& 0.7, & m_3(S) &=& 0.25
\end{array}
$$

Now, we obtain:

$$
\begin{array}{lcllcl}
(m_1 \circledcirc m_2 \circledcirc m_3)(\emptyset) &=& 1.25\cdot 10^{-4} & (m_1 \circledcirc m_2 \circledcirc m_3)(\{a,b,c\}) &=& 0.024\\
(m_1 \circledcirc m_2 \circledcirc m_3)(\{a,b\}) &=& 1.5\cdot 10^{-3} & (m_1 \circledcirc m_2 \circledcirc m_3)(\{b,c,d\}) &=& 0.028\\
(m_1 \circledcirc m_2 \circledcirc m_3)(\{b,c\}) &=& 0.002 & (m_1 \circledcirc m_2 \circledcirc m_3)(S) &=& 0.943\\
(m_1 \circledcirc m_2 \circledcirc m_3)(\{c,d\}) &=& 1.75\cdot 10^{-3} & & &
\end{array}
$$

By introducing small values for the empty set, we obtain non-zero values for the same sets than applying $\delta_{\mathcal{J}^{DS}}(u, \cdot)$. However, the combined values differ in several magnitudes, despite keeping the same level of certainty for $\{a, b\}$, $\{b, c\}$ and $\{c, d\}$ in the input. This is due to how these numbers are computed. For example, $\delta_{\mathcal{J}^{DS}}(u, \{b, c\}) = (1-p_1)\cdot p_2\cdot(1-p_3)$, while $(m_1 \circledcirc m_2 \circledcirc m_3)(\{b, c\}) = m_1(\emptyset)\cdot m_2(\{b, c\})\cdot m_3(\emptyset)$. As it can be observed, in the multi-layer case degrees of belief are computed by multiplying the certainty or uncertainty values of the available pieces of evidence. The degree of belief obtained by applying the disjunctive rule of combination, however, also considers the value of the empty set, which is interpreted as an estimation of the amount of conflict (Destercke and Burger, 2013). The multi-layer belief model accepts that degrees of belief must be based on justifications, but the empty set is never a justification. This means that the multi-layer belief model is not able to reproduce the disjunctive rule of combination. Nevertheless, $\mathrm{Bel}_{\mathcal{J}^{DS}}(u, \cdot)$ is another option to merge uncertain evidence in those contexts where the agent can only ensure that at least one of the sources is reliable. We leave the study of advantages and disadvantages of each option for future research.

Another frequent alternative to Dempster's rule of combination is the unnormalized rule of combination used in the transferable belief model (Definition 2.11). This setting allows mass functions to attach a positive value to the empty set. When computing belief based on these general mass functions, the value of the empty set will be added to the degree of belief of the total set, guaranteeing that the resulting function is a belief function without the need for normalization. The most natural attempt to replicate this rule by using the multi-layer belief model consists of considering the Dempster-Shafer frame of justification with a new allocation function t that we define in Proposition 3.10.

Proposition 3.10 *Given a set of evidence allocation functions $\mathfrak{F}$ and the function $t : 2^{\boldsymbol{\mathcal{E}}} \to \boldsymbol{\tau_{\mathcal{E}}}$ such that*

$$t(\boldsymbol{E}) = \begin{cases} \bigcap \boldsymbol{E} & \textit{if } \boldsymbol{E} \neq \emptyset, \textit{ and } \bigcap \boldsymbol{E} \neq \emptyset \\ S & \textit{otherwise.} \end{cases}$$

the set $\mathfrak{F} \cup \{t\}$ is a set of evidence allocation functions.

Proof.

To prove this, we will show that Definition 3.1 holds. By definition, we have that the empty set is mapped to S by every element of $\mathfrak{F} \cup \{t\}$. In addition, for those non-empty $\boldsymbol{E}$ such that $\bigcap \boldsymbol{E} \neq \emptyset$, $t(\boldsymbol{E}) = i(\boldsymbol{E})$, and $t(\boldsymbol{E}) = S$ otherwise. Therefore, $t(\boldsymbol{E})$ is a dense element of the topology $\boldsymbol{\tau_E}$—by Proposition 3.3—and $t(\boldsymbol{E}) \subseteq g(\boldsymbol{E})$ or $g(\boldsymbol{E}) \subseteq t(\boldsymbol{E})$ for every $g \in \mathfrak{F}$. □

We will call this function *transferable allocation function.* Proposition 3.11 demonstrates that the multi-layer belief model, set by the transferable allocation function and the Dempster-Shafer frame of justification, returns the same belief function as the transferable belief model.

Proposition 3.11 *Let $(S, \boldsymbol{\mathcal{E}}^Q)$ be a quantitative evidence frame, ℓ the number of elements in $\boldsymbol{\mathcal{E}}$, $t : 2^{\boldsymbol{\mathcal{E}}} \to \boldsymbol{\tau_{\mathcal{E}}}$ the transferable evidence allocation function defined in Proposition 3.10, and $\boldsymbol{\mathcal{J}}^{DS}$ Dempster-Shafer frame of justification. Let us consider the belief function $\mathrm{Bel}_{\boldsymbol{\mathcal{J}}^{DS}}(t, \cdot) : 2^S \to [0,1]$ defined by the multi-layer belief model and $bel : 2^S \to [0,1]$ a belief function obtained by applying the transferable belief model to the basic belief assignments $\{m_j | j = 1, \ldots, \ell\}$ such that $m_j(E_j) = p_j$, $m_j(S) = 1 - p_j$ for every $j = 1, \ldots, \ell$. Then,*

$$\mathrm{Bel}_{\boldsymbol{\mathcal{J}}^{DS}}(t, P) = bel(P)$$

for every $P \subseteq S$.

Proof.

Let us consider $m = \boxplus_j m_j$, where $\boxplus$ represents the unnormalized rule of combination. Given a proposition $P \subseteq S$, $\mathrm{Bel}_{\boldsymbol{\mathcal{J}}^{DS}}(t, P) = \sum_{A \subseteq P} \delta_{\boldsymbol{\mathcal{J}}^{DS}}(t, A)$. However, $\mathrm{bel}(P)$ is defined by cases, since the combined mass function m may give positive value to the empty set, and this value is not added to $\mathrm{bel}(P)$ when $P \neq \emptyset$ and $P \neq S$. In particular,

$$\text{bel}(P) = \begin{cases} 0 & \text{if } P = \emptyset, \\ \sum_{\emptyset \neq A \subseteq P} m(A) & \text{if } P \neq \emptyset \text{ and } P \neq S \\ 1 & \text{if } P = S. \end{cases}$$

Since Bel is a belief function, $\text{Bel}_{\boldsymbol{\mathcal{J}}^{DS}}(t, \emptyset) = 0$ and $\text{Bel}_{\boldsymbol{\mathcal{J}}^{DS}}(t, S) = 1$—for Proposition 3.7. So we need to show that

$$\sum_{A \subseteq P} \delta_{\boldsymbol{\mathcal{J}}^{DS}}(t, A) = \sum_{\emptyset \neq A \subseteq P} m(A)$$

for P such that $P \neq \emptyset$ and $P \neq S$. Since $\delta_{\boldsymbol{\mathcal{J}}^{DS}}(t, \emptyset) = 0$, this is equivalent to prove that $\delta_{\boldsymbol{\mathcal{J}}^{DS}}(t, A) = m(A)$ for $A \neq \emptyset$ and $A \neq S$. By Definition 2.11, we know that

$$m(A) = \sum_{\bigcap_j A_j = A} m_1(A_1) \cdot \ldots \cdot m_\ell(A_\ell),$$

exactly as the numerator of Equation (3.6). As $t(\boldsymbol{E}) = i(\boldsymbol{E})$ when $\bigcap \boldsymbol{E} \neq \emptyset$, the proof of Proposition 3.8 demonstrates that $m(A) = \delta_{\boldsymbol{\tau}}(t, A)$. Remembering that

$$\delta_{\boldsymbol{\mathcal{J}}^{DS}}(t, A) = \frac{\delta_{\boldsymbol{\tau}}(t, A)}{\sum_{T \in \boldsymbol{\mathcal{J}}^{DS}} \delta_{\boldsymbol{\tau}}(t, T)},$$

this means that showing $\sum_{T \in \boldsymbol{\mathcal{J}}^{DS}} \delta_{\boldsymbol{\tau}}(t, T) = 1$ would prove the result. Indeed, the empty set is the only element of $\boldsymbol{\tau}$ that is not contained in $\boldsymbol{\mathcal{J}}^{DS}$, so

$$\sum_{T \in \boldsymbol{\mathcal{J}}^{DS}} \delta_{\boldsymbol{\tau}}(t, T) = 1 - \delta_{\boldsymbol{\tau}}(t, \emptyset) = 1 - \sum_{\boldsymbol{E}: t(\boldsymbol{E}) = \emptyset} \delta(\boldsymbol{E}).$$

But $t(\boldsymbol{E})$ is never the empty set, so $\delta_{\boldsymbol{\mathcal{J}}^{DS}}(t, A) = \delta_{\boldsymbol{\tau}}(t, A)$ and the result holds. $\square$

This case differs to some extent from the representation of Dempster's rule of combination via the multi-layer belief model. When applying the multi-layer belief model with Dempster-Shafer frame of justification and evidence allocation function i, Dempster's rule of combination is replicated both at m (basic belief assignment) and bel (belief function) level. However, changing i for t replicates the belief function associated with the transferable belief model, but it does not replicate the unnormalized combination rule at the level of masses. This is because the general mass functions may assign a positive value to the empty set, while the empty set is never a justification, so $\delta_{\boldsymbol{\mathcal{J}}^{DS}}(f, \emptyset) = 0$ for every evidence allocation function f. If we accept not to allocate positive weights to the empty set, we have already defined an evidence allocation function that matches $\boldsymbol{E} \in 2^{\boldsymbol{\mathcal{E}}}$ with its intersection if it is not empty, and with something else otherwise: the function d defined in Proposition 3.5. When a collection of pieces of evidence $\boldsymbol{E} \in 2^{\boldsymbol{\mathcal{E}}}$ is such that $\bigcap \boldsymbol{E} \neq \emptyset$, then $i(\boldsymbol{E}) = t(\boldsymbol{E}) = d(\boldsymbol{E}) = \bigcap \boldsymbol{E}$. In contrast, when $\bigcap \boldsymbol{E} = \emptyset$, these three evidence allocation functions may return different images. First, $i(\boldsymbol{E}) = \emptyset$, which makes the normalization factor of Equation (3.4) different from 1. Second, $t(\boldsymbol{E}) = S$, which increases the value interpreted as uncertainty of the model. Lastly, $d(\boldsymbol{E})$ is the smallest dense set of the topology generated by $\boldsymbol{E}$ with $\bigcup \boldsymbol{E}$ as a total set. This means that $d(\boldsymbol{E})$ is never the empty set, so there is not normalization in $\delta_{\boldsymbol{\mathcal{J}}^{DS}}(d, \emptyset)$, and $d(\boldsymbol{E})$ is not necessarily the total set, which minimizes the increase of uncertainty of the model. In the next example, we compare the outcomes of $\mathrm{Bel}_{\boldsymbol{\mathcal{J}}^{DS}}(t, P)$, $\mathrm{Bel}_{\boldsymbol{\mathcal{J}}^{DS}}(d, P)$ and the unnormalized rule of combination to see these differences.

Example 3.4 Alternative to Unnormalized Rule of Combination

Let $S = \{a, b, c, d\}$ be a set of possible states and $\boldsymbol{\mathcal{E}}^Q = \{E_1, E_2, E_3\}$ where $E_1 = (\{a, b\}, 0.6)$, $E_2 = (\{b, c\}, 0.8)$ and $E_3 = (\{c, d\}, 0.7)$.

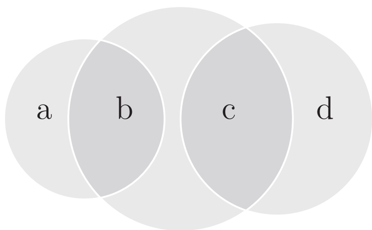


Figure 3.3: Representation of uncertain pieces of evidence by a Venn diagram. The size of the area represents the degree of certainty of the corresponding piece of evidence.

Fixing $\boldsymbol{\mathcal{J}}^{DS}$ as frame of justification and t as evidence allocation function, we obtain:

$$\begin{array}{rclrcl}
\delta_{\boldsymbol{\mathcal{J}}^{DS}}(t, \{a, b\}) & = & 0.036 & \delta_{\boldsymbol{\mathcal{J}}^{DS}}(t, \{b\}) & = & 0.144 \\
\delta_{\boldsymbol{\mathcal{J}}^{DS}}(t, \{b, c\}) & = & 0.096 & \delta_{\boldsymbol{\mathcal{J}}^{DS}}(t, \{c\}) & = & 0.224 \\
\delta_{\boldsymbol{\mathcal{J}}^{DS}}(t, \{c, d\}) & = & 0.056 & \delta_{\boldsymbol{\mathcal{J}}^{DS}}(t, S) & = & 0.444
\end{array}$$

Now, replacing t by d, we obtain:

$$\begin{array}{llll} \delta_{\mathcal{J}^{DS}}(d,\{a,b\}) = 0.036 & & \delta_{\mathcal{J}^{DS}}(d,\{b\}) = 0.144 \\ \delta_{\mathcal{J}^{DS}}(d,\{b,c\}) = 0.432 & & \delta_{\mathcal{J}^{DS}}(d,\{c\}) = 0.224 \\ \delta_{\mathcal{J}^{DS}}(d,\{c,d\}) = 0.056 & & \delta_{\mathcal{J}^{DS}}(d,S) = 0.108 \end{array}$$

Lastly, the unnormalized rule of combination returns these masses:

$$\begin{array}{llll} (m_1 \boxplus m_2 \boxplus m_3)(\emptyset) = 0.42 & & (m_1 \boxplus m_2 \boxplus m_3)(\{b\}) = 0.144 \\ (m_1 \boxplus m_2 \boxplus m_3)(\{a,b\}) = 0.036 & & (m_1 \boxplus m_2 \boxplus m_3)(\{c\}) = 0.224 \\ (m_1 \boxplus m_2 \boxplus m_3)(\{b,c\}) = 0.096 & & (m_1 \boxplus m_2 \boxplus m_3)(S) = 0.024 \\ (m_1 \boxplus m_2 \boxplus m_3)(\{c,d\}) = 0.056 & & \end{array}$$

As can be observed, while the transferable allocation function assigns a greater value to the total set (0.444) and the unnormalized rule of combination splits this value between the empty set (0.42) and the total set (0.024), the evidence allocation function d manages to assign a part of this weight to a proper subset of S, namely to $\{b,c\}$ ($\delta_{\mathcal{J}^{DS}}(d,\{b,c\}) = 0.432$ vs. $\delta_{\mathcal{J}^{DS}}(t,\{b,c\}) = (m_1 \boxplus m_2 \boxplus m_3)(\{b,c\}) = 0.096$).

These examples show how we can play with the different components of the multi-layer belief model to obtain some well-known belief models and reasonable alternatives. Although studying further cases is out of the scope of this thesis, by modifying other elements of the multi-layer belief model, we could replicate different families of belief models. For example, the family of belief functions for dependent sources (Denœux, 2008), whose combination rules are based on t-norms, could be replicated by defining a δ-function based on that t-norm.

We will further study the multi-layer belief model in Chapter 5, where we study the computational complexity of this evidence combination method among others. One open question that we leave for future research is the study of the mathematical properties of the multi-layer belief model. This includes defining combination and decombination of pieces of evidence within this context, and determining whether these operations are commutative, associative and have a unique neutral element. Another step towards the development of this model would be generalizing its definition for sources that provide multiple pieces of evidence. In Dempster-Shafer terminology, this means allowing basic belief assignments with more than two focal elements.

Throughout this chapter, we have defined a belief model that bridges Dempster-Shafer theory and topological models of evidence. Our analysis has primarily focused on the relationship between the multi-layer belief model and Dempster-Shafer theory. However, this model also has the potential to enrich topological models of evidence to handle uncertain evidence. This could be a valuable tool for research areas that intersects with topological models of evidence and uncertainty problems, such as argumentation (Baroni et al., 2018; Hung, 2017; Shi et al., 2017; Spaans and Doder, 2023), learning (Baltag et al., 2019a, 2019b; Gierasimczuk, 2010; Kelly and Lin, 2021; Vargas Sandoval, 2020), and reasoning under uncertainty (Corsi et al., 2023; Dubois et al., 2023; Flaminio et al., 2022; Halpern, 2017). In the next chapter, we initiate a discussion about logics for comparing strengths of beliefs based on uncertain evidence, aiming to define a logic for the multi-layer belief model.

3.3 Use Case Scenario: Navigation

The multi-layer belief model stands out for its adaptability compared to the belief models it is based on. By adaptability, we mean that the multi-layer belief model has two parameters—frame of justification and evidence allocation function—that allow it to merge evidence in different ways without needing a specific method for each. In contrast, Dempster's rule of combination and topological models of evidence produce an output fully determined by the evidence input. Therefore, using the multi-layer belief model is particularly relevant in cases where the same digital system will, e.g., be exposed to different situations that require varying levels of caution.

As an illustrative example, we propose the following situation: In 2022, a viral video on the internet showed how the screen of a car in autopilot got confused by the vehicle it was following: a horse-drawn carriage. On the screen, there was a figure in constant change: a truck, a pedestrian, a car, a pedestrian behind a car, an oncoming truck, etc. As mentioned, the system got confused. Navigation systems like this one are a good example where adaptability, in our sense, is crucial. The same car, with the same digital system, may consider different evidential demands—being more or less strict with evidence to generate belief—depending on whether it is in autopilot mode or assisting the driver.

Let us consider a concrete toy example. Suppose the sensors of the car system gather the following quantitative evidence set:

$$(\{\text{Human Silhouette}\}, 0.9)$$
$$(\{\text{Two Back Wheels}, \text{Large Height}\}, 0.65)$$
$$(\{\text{Two Back Wheels}, \text{Medium Width}\}, 0.7)$$

For simplicity, we denote Human Silhouette by HS, Two Back Wheels by $2BW$, Large Height by LH, and Medium Width by MW. So let $S = \{HS, 2BW, LH, MW\}$ be a set of possible states, and $\boldsymbol{\mathcal{E}}^Q = \{(\{HS\}, 0.9), (\{2BW, LH\}, 0.65), (\{2BW, MW\}, 0.7)\}$ be a quantitative evidence set.

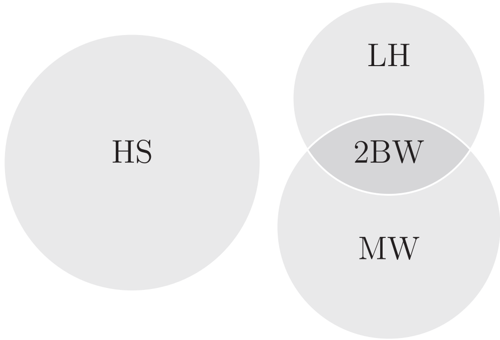


Figure 3.4: Representation of uncertain pieces of evidence in the running example by a Venn diagram. The size of the area represents the degree of certainty of the corresponding piece of evidence.

The feature *human silhouette* may be included in the internal ontology for pedestrian, motorbike, or bike, but not for truck and car. Similarly, the features *two back wheels* and *large height* together may belong to the ontology of truck, but not to car. Conversely, *two back wheels* and *medium width* together point to a car. Therefore, interpreting the gathered evidence all together is not straightforward.

Let us say that the car system uses the multi-layer belief model to identify what is in front of it. First, it may calculate the δ function (see Table 3.3), since it is independent of the parameter choice.

$$\delta(\emptyset) = 0.01$$

$$\delta(\{\{HS\}\}) = 0.09$$

$$\delta(\{\{2BW, LH\}\}) = 0.02$$

$$\delta(\{\{2BW, MW\}\}) = 0.02$$

$$\delta(\{\{HS\}, \{2BW, LH\}\}) = 0.18$$

$$\delta(\{\{HS\}, \{2BW, MW\}\}) = 0.22$$

$$\delta(\{\{2BW, LH\}, \{2BW, MW\}\}) = 0.05$$

$$\delta(\{\{HS\}, \{2BW, LH\}, \{2BW, MW\}\}) = 0.41$$

Table 3.3: Image of the δ function in the running example.

For illustrative purposes, let us explore different options for frames of justification and evidence allocation functions to place the results in context later on. In

Figure 3.5, we visually represent the elements of the Dempster-Shafer and strong denseness frames of justification.

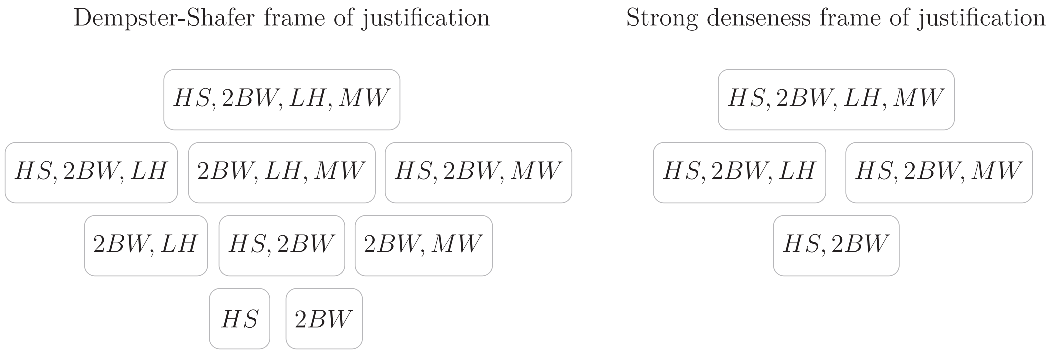


Figure 3.5: Visual representation of Dempster-Shafer and strong denseness frame of justification of the running example. Each box represents a set formed by the elements enclosed in it.

Table 3.4 collects the images of the evidence allocation functions i, u and d, along with the corresponding value of $\delta(\boldsymbol{E})$ (showing how δ values distribute after applying these evidence allocation functions).

$\boldsymbol{E}$	$\boldsymbol{i(E)}$	$\boldsymbol{u(E)}$	$\boldsymbol{d(E)}$	$\boldsymbol{\delta(E)}$
$\emptyset$	S	S	S	0.01
$\{HS\}$	$\{HS\}$	$\{HS\}$	$\{HS\}$	0.09
$\{2BW, LH\}$	$\{2BW, LH\}$	$\{2BW, LH\}$	$\{2BW, LH\}$	0.02
$\{2BW, MW\}$	$\{2BW, MW\}$	$\{2BW, MW\}$	$\{2BW, MW\}$	0.02
$\{HS\}, \{2BW, LH\}$	$\emptyset$	$\{HS, 2BW, LH\}$	$\{HS, 2BW, LH\}$	0.18
$\{HS\}, \{2BW, MW\}$	$\emptyset$	$\{HS, 2BW, MW\}$	$\{HS, 2BW, MW\}$	0.22
$\{2BW, LH\}, \{2BW, MW\}$	$\{2BW\}$	$\{2BW, LH, MW\}$	$\{2BW\}$	0.05
$\{HS\}, \{2BW, LH\}, \{2BW, MW\}$	$\emptyset$	S	$\{HS, 2BW\}$	0.41

Table 3.4: Images of δ by the evidence allocation functions $\{i, u, d\}$.

With this data in mind, let us compare three outputs of the multi-layer belief model: Dempster-Shafer frame of justification and intersection as evidence allocation function (DS + i) (the agent considers every argument and assumes all sources contain the correct element); Dempster-Shafer frame of justification and minimum dense set as evidence allocation function (DS + d) (the agent considers every argument and assumes a collection of sources contains the correct element); and strong denseness frame of justification and minimum dense set as evidence allocation function (SD + d) (the agent considers only arguments consistent with every other element of the evidential topology and assumes a collection of sources contains the correct element). We will compare the resulting $\delta_{\mathcal{J}}$ and belief functions with respect to the propositions: *There is a pedestrian in front of the car* ($\{HS\}$); *There is a truck in front of the car* ($\{2BW, LH\}$); *There is a car in front of the car* ($\{2BW, MW\}$); uncertainty of the model $\delta_{\mathcal{J}}(f, S)$; and proposition with highest $\delta_{\mathcal{J}}$ value.

Tables 3.5 and 3.6 display the results of $\mathrm{Bel}_{\mathcal{J}}(f, \cdot)$ and $\delta_{\mathcal{J}}(f, \cdot)$ respectively, for the corresponding frame of justification $\mathcal{J}$ and evidence allocation function f.

Proposition	P	$\mathrm{Bel}_{\mathcal{J}^{DS}}(i, P)$	$\mathrm{Bel}_{\mathcal{J}^{DS}}(d, P)$	$\mathrm{Bel}_{\mathcal{J}^{SD}}(d, P)$
Pedestrian	$\{HS\}$	0.47	0.09	0
Truck	$\{2BW, LH\}$	0.37	0.07	0
Car	$\{2BW, MW\}$	0.37	0.07	0

Table 3.5: Degrees of belief according to the multi-layer belief model.

$\delta_{\mathcal{J}}$	**Uncertainty** $\delta_{\mathcal{J}}(f, S)$	**Proposition with higher** $\delta_{\mathcal{J}}$ **value**	$\delta_{\mathcal{J}}(f, P)$
$\delta_{\mathcal{J}^{DS}}(i, \cdot)$	0.05	$\{HS\}$	0.47
$\delta_{\mathcal{J}^{DS}}(d, \cdot)$	0.01	$\{HS, 2BW\}$	0.41
$\delta_{\mathcal{J}^{SD}}(d, \cdot)$	0.12	$\{HS, 2BW\}$	0.5

Table 3.6: Values of $\delta_{\mathcal{J}}(f, \cdot)$ for corresponding $\mathcal{J}$ and f according to the multi-layer belief model.

One critical situation in this context is confusing a pedestrian for a motor vehicle or *vice versa*, as the required action may differ significantly. For instance, mistaking a pedestrian walking on the road could lead to honking and slowing down until stopping completely, potentially causing accidents if there is no actual pedestrian (thus unexpected behaviour for other drivers). However, encountering a truck ahead may simply require slowing down to maintain a safer distance. In an open space as a city road, such mistake is totally plausible. Depending on how the system is used (autonomous car, driven-assistance system, autopilot, etc.), there may be a preference to minimize false positives or negatives. In other words, one may prefer setting weaker evidential demands (Dempster-Shafer frame of justification) or stronger ones (strong denseness frame). Consider the scenario of a horse-drawn carriage: if the car detects a pedestrian and the protocol mandates honking and stopping, in a car with driver assistance—where the human driver observes the situation—honking could startle the horses and unexpectedly stopping may cause an accident. Conversely, in an autonomous car, failing to react to a potential pedestrian could be seen as negligent, possibly leading to an accident as well. In Tables 3.5 and 3.6, we observe that the belief model $\mathrm{Bel}_{\mathcal{J}^{SD}}(d, \cdot)$ adjusts to the situation of the car with driven-assistance, as it does not identify a pedestrian in this scenario regardless of the decision protocol ($\mathrm{Bel}_{\mathcal{J}^{SD}}(d, \{HS\}) = 0$). On the other hand, $\mathrm{Bel}_{\mathcal{J}^{DS}}(i, \cdot)$ or $\mathrm{Bel}_{\mathcal{J}^{DS}}(d, \cdot)$ may perform better in an autonomous car context, as they assign some degree of belief to the presence of a pedestrian ($\mathrm{Bel}_{\mathcal{J}^{DS}}(i, \{HS\}) = 0.47$, $\mathrm{Bel}_{\mathcal{J}^{DS}}(d, \{HS\}) = 0.09$), with the reaction depending on the decision protocol.

Another advantage of obtaining different belief results based on the same δ function is the potential for discovering unexpected outcomes through comparative exploration. In our example, we observe two models that assign higher weights to a proposition not matching any known ontology: $HS, 2BW$. This discrepancy could alert about the presence of an unknown object that intersects several ontologies (pedestrian, motorbike, car, truck, etc). Consequently, a suitable protocol might involve a cautious response encompassing actions common to these intersecting ontologies, such as slowing down and maintaining a safe distance without honking or stopping. Note that this serves merely as an illustrative example. It has been inspired by a chat with Dr. Philip Xu who works on information fusion for autonomous robots—some of his work related to Dempster-Shafer theory in (Tong et al., 2019, 2021). Other scenarios where a digital system operates in varied contexts, with specific preferences about avoiding false negatives or false positives, may include software for medical diagnostics or legal support.

Chapter 4

A Qualitative Logic for Evidence and Belief Comparison

In the previous chapter, we proposed a so-called *multi-layer belief model* for measuring degrees of beliefs for an agent who possesses a body of possibly *mutually contradictory, partial,* and *uncertain* pieces of evidence and whose *evidential demands* may change depending on the context. Now, we take a first step towards developing a modal logic for this belief model. In particular, in this chapter we propose a qualitative logic for explicitly comparing (1) strengths of beliefs and (2) evidential support with respect to degrees of certainty about evidence. To the best of our knowledge, (2) is the novel component of the logic and formalized via an *order-lifting* of a total pre-order on a set of uncertain evidence pieces. This component is thus intended to compare propositions with respect to the degrees of certainty of the *sets* of evidence pieces—rather than of individual evidence pieces—supporting them. Therefore, our logic can be seen as a qualitative logic for evidence and belief comparison.

The chapter is structured as follows. In Section 4.1, we discuss some works that inspire and motivate ours. In Section 4.2, we introduce and justify an order lifting for preference orders that preserves good properties when applied on uncertain evidence. Section 4.3 presents our proposal for a qualitative logic to compare strengths of evidence in terms of certainty and belief. We conclude the chapter with Section 4.4, in which we present a real-world scenario that illustrates the importance of reasoning about uncertain evidence and degrees of belief within the same logical system.

4.1 Context

Throughout this chapter, we will adopt the evidence representation established in Section 2.3. Therefore, we assume that S is a finite set of possible states, $\boldsymbol{\mathcal{E}}^Q$ is a quantitative evidence set, and $(S, \boldsymbol{\mathcal{E}}^Q)$ is a quantitative evidence frame. In addition, we assume familiarity with the terms introduced in Chapter 3. In particular, we will denote as $\mathrm{Bel}(\cdot)$ the belief function obtained by applying the multi-layer belief model with the Dempster-Shafer frame of justification and the intersection evidence allocation function (denoted by $\mathrm{Bel}_{\mathcal{J}^{DS}}(i, \cdot)$ in Chapter 3). This belief function is exactly the belief function obtained by applying Dempster's rule of combination to a collection of simple basic belief assignments determined by $\boldsymbol{\mathcal{E}}^Q$ (see Proposition 3.8).

The goal of this chapter is to define a qualitative logic that compares the strength of belief and evidential support with respect to the certainty values of the available evidence. The former is inspired by (Harmanec and Hájek, 1994). This work presents a qualitative belief logic (QBL) for comparing strengths of belief, where belief is represented by a belief function (see Definition 2.5), but not necessarily a belief function generated by combining specific pieces of evidence by a specific combination method. We aim to use the logic of Harmanec and Hájek (1994) in this respect. That is, we aim to provide a logic that reasons about the strength of belief not only according to a given belief function, but according to the belief function obtained by a given set of quantitative evidence. To this end, we equip our formal models with $\boldsymbol{\mathcal{E}}^Q$ and Bel, defined as in the previous paragraph. As a formal language, we use two binary modal operators. One to compare propositions with respect to their degrees of belief (equivalent to the one introduced in (Harmanec and Hájek, 1994)); and one to compare propositions with respect to the degrees of certainty of the evidence supporting them. Defining this second modal operator is the challenge we address in this chapter. Consequently, we will thoroughly discuss orders on $\boldsymbol{\mathcal{E}}$, $2^{\boldsymbol{\mathcal{E}}}$ and 2^S. For clarity, the reader can keep in mind that we will use the following notation: Given a quantitative evidence frame $(S, \boldsymbol{\mathcal{E}}^Q)$, we denote by $(Q, \leq)$ the order among certainty values that follows the real order in the interval $(0, 1)$, by $(\boldsymbol{\mathcal{E}}, \leq_e)$ the order among evidence pieces derived from $(Q, \leq)$, by $(2^{\boldsymbol{\mathcal{E}}}, \preceq_e)$ any order lifting from $(\boldsymbol{\mathcal{E}}, \leq_e)$, and by $(2^S, \trianglelefteq_e)$ the order among propositions that we are aiming at and that will be based on the previous ones.

Beyond the logic of Harmanec and Hájek (1994), there is a rather long tradition linking modal logic and Dempster-Shafer theory of belief functions, as well as modal logic and topological reasoning—the other cornerstone of the multi-layer belief model. Our approach in this paper falls under both traditions and connects to many such logics thereof. One of the most relevant works in modal logic that understands evidence as possibly partial and mutually contradictory pieces of information is (Van Benthem and Pacuit, 2011). This work has been developed further in (Van Benthem et al., 2012, 2014) and inspired the already discussed topological models of evidence of (Baltag et al., 2022). This way of modeling evidence, however, does not take into account the degrees of (un)certainty about evidence: evidence possession is modelled in a binary manner; agents either have or do not have evidence that support propositions. Modal logics that reason about uncertain evidence or, specifically, about degrees of belief based on uncertain evidence, can be found among the logics for belief functions. Starting with (Ruspini, 1987), this line of research has tended to capture belief functions in epistemic logic, resulting in multi-valued logics that are capable of representing graded belief. A recent proposal is (Dubois et al., 2023), where we can find a concrete multi-valued logic for belief and an overview of the most relevant publications in this direction. Furthermore, if we look at other frameworks related to belief and uncertainty, we find other comparison logics, such as those defined in (Ghosh and De Jongh, 2013), which explicitly reason about different strengths of belief, and (Ding et al., 2021), which presents a family of comparison logics for the uncertainty framework of imprecise probabilities.

In our case, we chose the logic of Harmanec and Hájek (1994) (QBL) as a starting point because its qualitative nature facilitates its connection to topological models of evidence, a relevant property for the long-term goal of developing a modal logic for the multi-layer belief model. As mentioned above, in this chapter we extend QBL to reason not only about the strength of beliefs, but also about the certainty degrees of evidential support. This logic contributes to the relevant literature in two different ways. First, it opens the door to extending topological models of evidence to a quantitative version that captures different degrees of certainty or preferences among pieces of evidence. Second, it extends QBL by providing a representation of basic evidence pieces that generate degrees of belief through Dempster's combination rule.

4.2 An Order Lifting for Uncertain Evidence

In our logic—to be formally defined in Section 4.3—we compare propositions with respect to the degree of certainty of the evidential support they receive. It is easy to compare basic pieces of evidence with respect to their degrees of uncertainty: we simply order pieces of evidence according to certainty values specified in $\boldsymbol{\mathcal{E}}^Q$. That is, given a quantitative evidence frame $(S, \boldsymbol{\mathcal{E}}^Q)$ and pairs $(E, q_E), (E', q_{E'}) \in \boldsymbol{\mathcal{E}}^Q$, we define an order $\leq_e$ on $\boldsymbol{\mathcal{E}}$ as follows:

$$E \leq_e E' \text{ if and only if } q_E \leq q_{E'} \qquad \text{(order } \leq_e\text{)}$$

where $\leq$ is the standard total pre-order defined on $(0, 1)$. We say E' *certainty-dominates* E when $E \leq_e E'$. When $q_E < q_{E'}$, we say E' *strictly certainty-*

dominates E and denote it by $E <_e E'$. "Certainty-dominates" simply means "at least as certain as", we prefer the former reading as the latter leads to convoluted readings of the comparison operators we later define over sets of propositions. It is not difficult to see that $(\boldsymbol{\mathcal{E}}, \leq_e)$ forms a total pre-order ($\leq_e$ is not necessarily a partial order on $\boldsymbol{\mathcal{E}}$ since it could be that $q_E = q_{E'}$ but $E \neq E'$).

In order to compare propositions with respect to the degree of certainty of the evidential support they receive, we need to define an order on *sets* of pieces of evidence that support propositions. To this end, we need to *extend* or *lift* the order $\leq_e$ to an order $\preceq_e$ on $2^{\boldsymbol{\mathcal{E}}}$. We can then easily extend this order to an order on 2^S by simply linking sets of pieces of evidence to the propositions they support.

4.2.1 Egli-Milner Order for Uncertainty

In (Van Benthem et al., 2009), the authors present four immediate ways of lifting a pre-order: ∀∀-, ∀∃-, ∃∀-, and ∃∃-liftings. We start our discussion by analyzing the implications of these liftings of $\leq_e$ in our specific context. Our goal is to identify the one that better captures our intention of comparing propositions based (only) on the degree of certainty of their evidential support. Recall that we assume a finite set of pieces of evidence $\boldsymbol{\mathcal{E}}$ defined on a finite set of possible states S. In addition, we represent degree of certainty of evidence as values within the interval $(0, 1)$, so they are linearly ordered and the corresponding pair $(\boldsymbol{\mathcal{E}}, \leq_e)$ is a total pre-order, that is, a total, reflexive and transitive order—see, e.g., (Roman, 2008) for such standard definitions of order theory.

When defining a lifting $(2^{\boldsymbol{\mathcal{E}}}, \preceq_e)$ for the total pre-order $(\boldsymbol{\mathcal{E}}, \leq_e)$, the following definitions will be useful. For any $\boldsymbol{E}, \boldsymbol{E}' \in \boldsymbol{\mathcal{E}}$:

1. If $\boldsymbol{E} \preceq_e \boldsymbol{E}'$ and $\boldsymbol{E}' \preceq_e \boldsymbol{E}$, then $\boldsymbol{E} \equiv_e \boldsymbol{E}'$ and we say $\boldsymbol{E}$ *and* $\boldsymbol{E}'$ *are equally certain.*
2. If $\boldsymbol{E} \preceq_e \boldsymbol{E}'$ and $\boldsymbol{E}' \not\preceq_e \boldsymbol{E}$, then $\boldsymbol{E} \prec_e \boldsymbol{E}'$ and we say $\boldsymbol{E}'$ *strictly certainty-dominates* $\boldsymbol{E}$.
3. If $\boldsymbol{E} \not\preceq_e \boldsymbol{E}'$ and $\boldsymbol{E}' \not\preceq_e \boldsymbol{E}$, then $\boldsymbol{E} ?_e \boldsymbol{E}'$ and we say $\boldsymbol{E}$ *and* $\boldsymbol{E}'$ *are incomparable.*

Additionally, when we have chosen the order lifting that best suits our purposes, we will place the empty set at the bottom of the obtained order. So we will discuss the four liftings presented in (Van Benthem et al., 2009) by excluding the empty set.

∀∀-lifting If we define an evidence comparison order $\preceq_e$ on $2^{\boldsymbol{\mathcal{E}}} \setminus \{\emptyset\}$ as

$$\boldsymbol{E} \preceq_e \boldsymbol{E}' \text{ if and only if } \forall E \in \boldsymbol{E}\ \forall E' \in \boldsymbol{E}' : E \leq_e E',$$

then it is not possible to compare some sets of evidence whose certainty-domination is intuitively clear. In Figure 4.1 we visualize one of these cases. We observe that $\forall\forall$-lifting gives the intuitive result in case (a): $\boldsymbol{E'}$ strictly certainty-dominates $\boldsymbol{E}$. However, in case (b), $\forall\forall$-lifting says that $\boldsymbol{E}$ and $\boldsymbol{E'}$ are incomparable. Ideally, we would like to be able to compare these not extreme cases of $\boldsymbol{E}$ and $\boldsymbol{E'}$ depicted in Figure 4.1.(b) and conclude that $\boldsymbol{E} \preceq_e \boldsymbol{E}'$, so this order lifting is too conservative for our purposes.

(a) $\boldsymbol{E} \prec_e \boldsymbol{E}'$ w.r.t. $\forall\forall$-lifting. (b) $\boldsymbol{E} ?_e \boldsymbol{E}'$ w.r.t. $\forall\forall$-lifting.

Figure 4.1: Brown segments represent the range of certainty values of the pieces of evidence in $\boldsymbol{E}$, so the vertical lines in the extremes represent the minimum and maximum of these values. Purple segments represent the respective values for $\boldsymbol{E}'$.

$\exists\exists$-lifting If we define an evidence comparison order $\preceq_e$ on $2^{\boldsymbol{\mathcal{E}}} \setminus \{\emptyset\}$ as

$$\boldsymbol{E} \preceq_e \boldsymbol{E'} \text{ if and only if } \exists E \in \boldsymbol{E}\ \exists E' \in \boldsymbol{E'} : E \leq_e E',$$

then we would again obtain a counter-intuitive comparisons for the case depicted in Figure 4.1.(b). As noted in Figure 4.2.(b), $\exists\exists$-lifting establishes $\boldsymbol{E} \equiv_e \boldsymbol{E}'$ in this case, which contradicts intuitions since every piece of evidence in $\boldsymbol{E}$ is strictly certainty-dominated by some evidence in $\boldsymbol{E}'$.

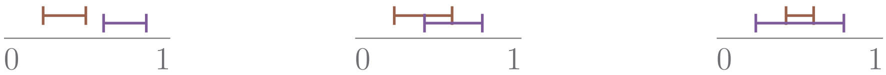
(a) $\boldsymbol{E} \prec_e \boldsymbol{E}'$ w.r.t. $\exists\exists$-lifting. (b) $\boldsymbol{E} \equiv_e \boldsymbol{E}'$ w.r.t. $\exists\exists$-lifting.

Figure 4.2: Brown segments represent the range of certainty values of the pieces of evidence that support $\boldsymbol{E}$, so the vertical lines in the extremes represent the minimum and maximum of these values. Purple segments represent the respective values for $\boldsymbol{E}'$.

$\forall\exists$-lifting If we define an evidence comparison order $\preceq_e$ on $2^{\boldsymbol{\mathcal{E}}} \setminus \{\emptyset\}$ as

$$\boldsymbol{E} \preceq_e \boldsymbol{E'} \text{ if and only if } \forall E \in \boldsymbol{E}\ \exists E' \in \boldsymbol{E'} : E \leq_e E',$$

then we would again obtain counter-intuitive comparisons for some cases. In Figure 4.3, we observe that this order lifting captures intuitions for cases (a) and (b), but its answer for case (c) is not the expected one. In this case, $\forall\exists$-lifting says that $\boldsymbol{E}'$ strictly certainty-dominates $\boldsymbol{E}$. However, $\boldsymbol{E}$ strictly certainty-dominates some evidence in $\boldsymbol{E}'$ as well, so stating $\boldsymbol{E} \prec_e \boldsymbol{E}'$ contradicts intuitions.



(a) $\boldsymbol{E} \prec_e \boldsymbol{E}'$ w.r.t. $\forall\exists$-lifting.

(b) $\boldsymbol{E} \prec_e \boldsymbol{E}'$ w.r.t. $\forall\exists$-lifting.

(c) $\boldsymbol{E} \prec_e \boldsymbol{E}'$ w.r.t. $\forall\exists$-lifting.

Figure 4.3: Brown segments represent the range of certainty values of the pieces of evidence that support $\boldsymbol{E}$, so the vertical lines in the extremes represent the minimum and maximum of these values. Purple segments represent the respective values for $\boldsymbol{E}'$.

$\exists\forall$-lifting If we define an evidence comparison order $\preceq_e$ on $2^{\boldsymbol{\mathcal{E}}} \setminus \{\emptyset\}$ as

$$\boldsymbol{E} \preceq_e \boldsymbol{E'} \text{ if and only if } \exists E \in \boldsymbol{E}\ \forall E' \in \boldsymbol{E'} : E \leq_e E',$$

then we are in a similar situation as in Figure 4.3 but concluding $\boldsymbol{E}' \prec_e \boldsymbol{E}$ instead of $\boldsymbol{E} \prec_e \boldsymbol{E}'$ in case (c). We draw similar conclusions.

Bi-directional liftings. All $\forall\forall$-, $\forall\exists$-, $\exists\forall$-, $\exists\exists$-liftings present some counter-intuitive behavior in our context. In $\forall\forall$- and $\exists\exists$-liftings, the problem is intrinsic to their definition. However, in $\forall\exists$- and $\exists\forall$-liftings, problems arise from the lack of bi-directionallity. This can be solved by considering their bi-directional version, that is, by combining these liftings with their reverse.

Definition 4.1 ($\forall\exists$ bi-directional lifting). Let $(S, \boldsymbol{\mathcal{E}}^Q)$ be a quantitative evidence frame. We define the $\forall\exists$ bi-directional lifting of $(\boldsymbol{\mathcal{E}}, \leq_e)$ as

$$\boldsymbol{E} \preceq_e \boldsymbol{E'} \quad \text{if and only if} \quad \begin{array}{l} (1)\ \forall E \in \boldsymbol{E}\ \exists E' \in \boldsymbol{E'} : E \leq_e E';\ \text{and} \\ (2)\ \forall E' \in \boldsymbol{E'}\ \exists E \in \boldsymbol{E} : E \leq_e E'. \end{array}$$

The order $(2^{\boldsymbol{\mathcal{E}}}, \preceq_e)$ is called *Egli-Milner order*. In (Shi and Sun, 2021), the authors study the logic of this order as a preference lifting. The authors of this paper remark other uses of the Egli-Milner order in the literature, such as works on domain theory (Plotkin, 1976), verisimilitude (Van Benthem, 1987; Brink, 1992) or truthmaker semantics (Fine, 2017).

Definition 4.2 ($\exists\forall$ bi-directional lifting). Let $(S, \boldsymbol{\mathcal{E}}^Q)$ be a quantitative evidence frame. We define the $\exists\forall$ bi-directional lifting of $(\boldsymbol{\mathcal{E}}, \leq_e)$ as

$$\boldsymbol{E} \preceq_e \boldsymbol{E'} \quad \text{if and only if} \quad \begin{array}{l} (1)\ \exists E \in \boldsymbol{E}\ \forall E' \in \boldsymbol{E'} : E \leq_e E';\ \text{and} \\ (2)\ \exists E' \in \boldsymbol{E'}\ \forall E \in \boldsymbol{E} : E \leq_e E'. \end{array}$$

In our particular context, where the evidence set $\boldsymbol{\mathcal{E}}$ is finite and $\leq_e$ is a total pre-order, $\forall\exists$ bi-directional lifting and $\exists\forall$ bi-directional lifting are equivalent.

Proposition 4.1 *Let $(S, \boldsymbol{\mathcal{E}}^Q)$ be a quantitative evidence set such that S and $\boldsymbol{\mathcal{E}}$ are finite. Then, for all $\boldsymbol{E}, \boldsymbol{E}' \in 2^{\boldsymbol{\mathcal{E}}} \setminus \{\emptyset\}$:*

$$\begin{array}{l} (1)\ \forall E \in \boldsymbol{E}\ \exists E' \in \boldsymbol{E'} : E \leq_e E';\ \textit{and} \\ (2)\ \forall E' \in \boldsymbol{E'}\ \exists E \in \boldsymbol{E} : E \leq_e E' \end{array}$$

if and only if

$$\begin{array}{l} (1')\ \exists E \in \boldsymbol{E}\ \forall E' \in \boldsymbol{E'} : E \leq_e E';\ \textit{and} \\ (2')\ \exists E' \in \boldsymbol{E'}\ \forall E \in \boldsymbol{E} : E \leq_e E'. \end{array}$$

Proof.

Right-left direction is true in general, since $(1')$ implies (2) and $(2')$ implies (1). To prove left-right direction, suppose that (1) and (2) are the case. Since $(\boldsymbol{\mathcal{E}}, \leq_e)$ is a total pre-order on a finite set and, thus, $\boldsymbol{E}$ is finite, there is a piece of evidence $E_{max} \in \boldsymbol{E}$ such that $E \leq_e E_{max}$ for all $E \in \boldsymbol{E}$. By (1), we also have that there is $E' \in \boldsymbol{E}'$ such that $E_{max} \leq_e E'$. As $\leq_e$ is transitive, we obtain that $E \leq_e E'$ for all $E \in \boldsymbol{E}$. Hence, $(2')$ is satisfied. Similarly, the finiteness and totality of $(\boldsymbol{\mathcal{E}}, \leq_e)$ also implies that there exists $E_{min} \in \boldsymbol{E}$ such that $E_{min} \leq_e E$ for all $E \in \boldsymbol{E}$. By (2), $E_{min} \leq_e E'$ for all $E' \in \boldsymbol{E'}$, hence $(1')$ is also satisfied. □

In contexts where uncertainty is not represented by a total pre-order over a finite space, further discussion is needed to differentiate these two order liftings and their implications. In our case, we can use both definitions indistinguishably for Proposition 4.1. We will stick to $\forall\exists$ bi-directional lifting (or Egli-Milner order) for the connections of this work with (Shi and Sun, 2021). In Definition 4.3, we present our proposal for a certainty order among sets of pieces of evidence.

Definition 4.3 (Evidence Comparison Order $(2^{\boldsymbol{\mathcal{E}}}, \preceq_e)$). Given a finite quantitative evidence frame $(S, \boldsymbol{\mathcal{E}}^Q)$, the *evidence comparison order* $\preceq_e$ on $2^{\boldsymbol{\mathcal{E}}}$ is defined as follows. For all $\boldsymbol{E}, \boldsymbol{E'} \in 2^{\boldsymbol{\mathcal{E}}}$:

1. $\emptyset \preceq_e \boldsymbol{E}$, and
2. for all $\boldsymbol{E} \neq \emptyset$:

$$\boldsymbol{E} \preceq_e \boldsymbol{E'} \quad \text{if and only if} \quad \begin{array}{l}(1)\ \forall E \in \boldsymbol{E}\ \exists E' \in \boldsymbol{E'} : E \leq_e E';\ \text{and} \\ (2)\ \forall E' \in \boldsymbol{E'}\ \exists E \in \boldsymbol{E} : E \leq_e E'.\end{array}$$

Observe that, restricted to non-empty sets of evidence pieces, $\preceq_e$ is the Egli-Milner extension/lifting of $\leq_e$ on $\boldsymbol{\mathcal{E}}$. Given any $\boldsymbol{E}, \boldsymbol{E'} \in 2^{\boldsymbol{\mathcal{E}}}$: $\boldsymbol{E} \preceq_e \boldsymbol{E'}$ says that the evidence set $\boldsymbol{E'}$ is *at least as certain as* $\boldsymbol{E}$ in the following sense: (1) every piece of evidence in $\boldsymbol{E}$ is certainty-dominated by some piece of evidence in $\boldsymbol{E'}$, and (2) every piece of evidence in $\boldsymbol{E'}$ certainty-dominates some piece of evidence in $\boldsymbol{E}$, with the caveat that the empty set is always the least certain. We stipulate that the empty set is comparable according to the ordering $\preceq_e$ to enable our logic to distinguish between propositions supported by some evidence and those that are not. This creates a slight abuse of language, as the empty set is not truly certainty-dominated by every set of evidence, but rather 'evidence'-dominated by every set of evidence, in the sense that any set $\boldsymbol{E} \in 2^{\boldsymbol{\mathcal{E}}} \setminus \{\emptyset\}$ contains more pieces of evidence than the empty set. Therefore, every non-empty set of evidence dominates the empty set and, between two non-empty sets of evidence, domination depends on the certainty degree of the available evidence.

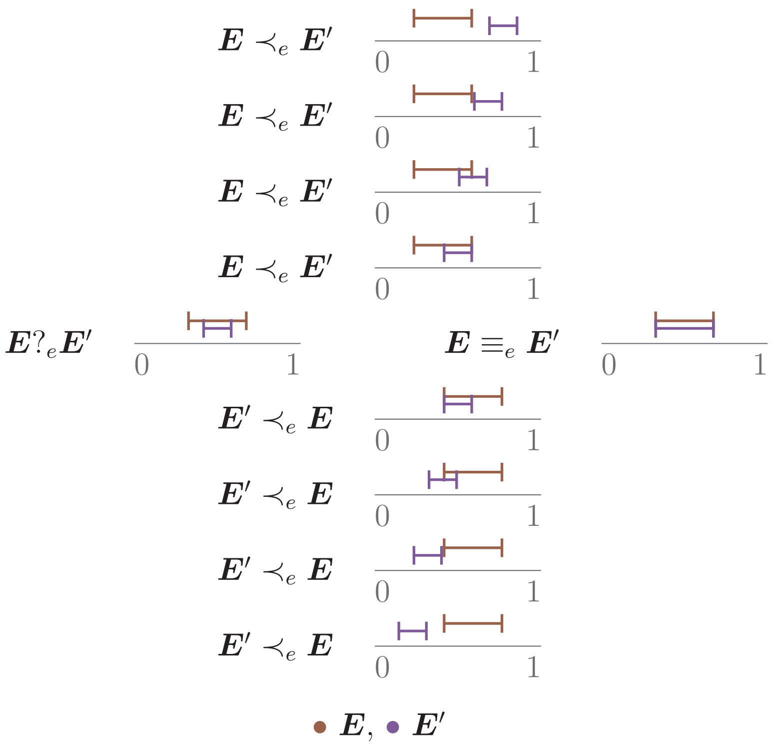


Figure 4.4: Visualization of $\preceq_e$. Brown segments represent the range of certainty values of the pieces of evidence that support $\boldsymbol{E}$, so the vertical lines in the extremes represent the minimum and maximum of these values. Purple segments represent the respective values for $\boldsymbol{E}'$.

We complete this definition by reminding the reader the shortcuts defined at the beginning of this section. That is, $\boldsymbol{E} \equiv_e \boldsymbol{E}'$ if and only if $\boldsymbol{E} \preceq_e \boldsymbol{E}'$ and $\boldsymbol{E}' \preceq_e \boldsymbol{E}$; $\boldsymbol{E} \prec_e \boldsymbol{E}'$ if and only if $\boldsymbol{E} \preceq_e \boldsymbol{E}'$ and $\boldsymbol{E}' \not\preceq_e \boldsymbol{E}$; and $\boldsymbol{E} ?_e \boldsymbol{E}'$ if and only if $\boldsymbol{E} \not\preceq_e \boldsymbol{E}'$ and $\boldsymbol{E}' \not\preceq_e \boldsymbol{E}$. In Figure 4.4 we visually represent the behaviour of the ordering $\preceq_e$.

4.2.2 Properties and Suitability

As an order on sets of pieces of evidence that compares sets of evidence pieces according to their degrees of certainty, we proposed the lifting of $(\boldsymbol{\mathcal{E}}, \leq_e)$ given in Definition 4.3. Now, we study the properties of this order. We also discuss how appropriate these properties are in our context of uncertain evidence and belief functions, justifying our choice.

Ordering properties

Proposition 4.2 *Given a quantitative evidence frame $(S, \boldsymbol{\mathcal{E}}^Q)$, the ordered pair $(2^{\boldsymbol{\mathcal{E}}}, \preceq_e)$ defined as in Definition 4.3, and $\boldsymbol{E}, \boldsymbol{E}' \in 2^{\boldsymbol{\mathcal{E}}}$, we have:*

1. *The pair $(2^{\boldsymbol{\mathcal{E}}}, \preceq_e)$ is a pre-order which is not necessarily total.*
2. *If $\boldsymbol{E} \preceq_e \emptyset$, then $\boldsymbol{E} = \emptyset$.*
3. *$\emptyset$ is the unique minimal element of $(2^{\boldsymbol{\mathcal{E}}}, \preceq_e)$.*
4. *$\boldsymbol{E} \subseteq \boldsymbol{E}'$ does not imply $\boldsymbol{E} \preceq_e \boldsymbol{E}'$ (i.e., $\preceq_e$ is not monotonic).*

Proof.

1. It follows immediately by the definition of $\preceq_e$ that it is reflexive and transitive. To see that $\preceq_e$ is not total, consider $S = \{a, b, c\}$ and $\boldsymbol{\mathcal{E}} = \Big\{(\{a\}, 0.2), (\{b\}, 0.1)(\{c\}, 0.3)\Big\}$, $\boldsymbol{E} = \Big\{\{a\}\Big\}$ and $\boldsymbol{E}' = \Big\{\{b\}, \{c\}\Big\}$. We then have $\boldsymbol{E} \not\preceq_e \boldsymbol{E}'$ since $\{b\} \in \boldsymbol{E}'$ does not certainty-dominate any element in $\boldsymbol{E}$, violating Definition 4.3.2.(2). Moreover, $\boldsymbol{E}' \not\preceq_e \boldsymbol{E}$ since $\{c\} \in \boldsymbol{E}'$ is not certainty-dominated by any element in $\boldsymbol{E}$, violating Definition 4.3.2.(1).
2. Suppose, toward contradiction, that $\boldsymbol{E} \preceq_e \emptyset$ and $\boldsymbol{E} \neq \emptyset$. The former, by Definition 4.3.2.(1), means that for all $E \in \boldsymbol{E}$, there is $E' \in \emptyset$ such that $E \leq_e E'$, which cannot be the case as the empty set has no elements. Therefore, if $\boldsymbol{E} \preceq_e \emptyset$, then $\boldsymbol{E} = \emptyset$.
3. Item 2 and Definition 4.3.1 together guarantee that $\boldsymbol{E} \preceq_e \emptyset$ if and only if $\boldsymbol{E} = \emptyset$, thus, $\emptyset$ is the unique minimal element of $(2^{\boldsymbol{\mathcal{E}}}, \preceq_e)$.

4. Consider $\boldsymbol{\mathcal{E}}^Q = \{(\{a\}, 0.2), (\{b\}, 0.3), (\{c\}, 0.1)\}$, $\boldsymbol{E} = \{\{a\}\}$, and $\boldsymbol{E'} = \{\{a\}, \{c\}\}$. In this case, $\boldsymbol{E} \subseteq \boldsymbol{E}'$ but $\boldsymbol{E} \not\preceq_e \boldsymbol{E}'$, since $\{c\} \in \boldsymbol{E'}$ violates Definition 4.3.2.(2): it does not certainty-dominates any evidence in $\boldsymbol{E}$.

□

From all these properties, we remark the fact of having the empty set as the minimum of $(2^{\boldsymbol{\mathcal{E}}}, \preceq_e)$. This implies that (1) the lack of evidence will not dominate any set of pieces of evidence according to our ordering, and (2) we can identify the propositions that are not supported by any evidence. Without Definition 4.3.1, every $\boldsymbol{E} \in 2^{\boldsymbol{\mathcal{E}}} \setminus \{\emptyset\}$ would be incomparable to $\emptyset$. However, thanks to our current definition of $\preceq_e$, we can identify those propositions that are not supported by the available evidence by Proposition 4.2.2.

We also highlight the non-monotonicity property described in Proposition 4.2.4. According to the set representation of uncertainty used in Dempster-Shafer theory, the order $(\boldsymbol{\mathcal{E}}, \leq_e)$ is non-monotonic. For example, given $S = \{a, b, c\}$ and $\boldsymbol{\mathcal{E}} = \Big\{(\{a\}, 0.7), (\{a, b\}, 0.2)\Big\}$, $\{a\} \subseteq \{a, b\}$ and $\{a, b\} \leq_e \{a\}$. Therefore, non-monotonicity for the order resulting from the lifting may be naturally accepted in this context.

Order lifting properties So far, intuitions and order properties of $\preceq_e$ are convenient for our context. Now, we discuss how good $(2^{\boldsymbol{\mathcal{E}}}, \preceq_e)$ is as an order lifting of $(\boldsymbol{\mathcal{E}}, \leq_e)$, and if we could get any better alternative. The classic works of preference lifting present the following properties as desirable properties for an order lifting (Barberà et al., 2004). In this part, we will use *order extension* as a synonym of *order lifting* to match the terminology of the cited sources.

Definition 4.4 (Order Lifting Properties). Let $(X, \leq)$ be a total order over a finite set and $(2^X, \preceq)$ an order lifting. We define the following properties.

(*Extension rule*) For all x_1, $x_2 \in X$, if $x_1 < x_2$ then $\{x_1\} \prec \{x_2\}$.

(*Extension Dominance*) For all $Y \subseteq X$ and all $x \in X \setminus Y$, (1) if $x < y$ for all $y \in Y$ then $Y \cup \{x\} \prec Y$, and (2) if $y < x$ for all $y \in Y$ then $Y \prec Y \cup \{x\}$.

(*Extension Independence*) For all Y_1, $Y_2 \subseteq X$ and $x \in X \setminus \{Y_1 \cup Y_2\}$, if $Y_1 \prec Y_2$ then $Y_1 \cup \{x\} \preceq Y_2 \cup \{x\}$.

(*Extension Strict Independence*) For all Y_1, $Y_2 \subseteq X$ and $x \in X \setminus \{Y_1 \cup Y_2\}$, if $Y_1 \prec Y_2$ then $Y_1 \cup \{x\} \prec Y_2 \cup \{x\}$.

If we consider the restriction of the order $(2^{\boldsymbol{\mathcal{E}}}, \preceq_e)$ to $(2^{\boldsymbol{\mathcal{E}}} \setminus \{\emptyset\}, \preceq_e)$ (so the Egli-Milner order), we obtain an order lifting of $(\boldsymbol{\mathcal{E}}, \leq_e)$ equivalent to the order lifting *min-max* defined in the context of preference lifting in Social Choice (Maly, 2020).

Definition 4.5 (Min-max Extension). Given a total order $(X, \leq)$ over a finite set and Y_1, $Y_2 \subseteq X$, we define the *min-max extension* $(2^X, \preceq)$ as

$$Y_1 \preceq Y_2 \quad \text{if and only if} \quad \begin{array}{l} \max(Y_1) \leq \max(Y_2) \text{ and} \\ \min(Y_1) \leq \min(Y_2) \end{array}$$

where $\max(Y_i) \in Y_i$ and $\min(Y_i) \in Y_i$ such that $y' \leq \max(Y_i)$ and $\min(Y_i) \leq y'$ for all $y' \in Y_i$, $i \in \{1, 2\}$. For any $i \in \{1, 2\}$, both elements $\max(Y_i)$ and $\min(Y_i)$ exists for Y_i finite.

Proposition 4.3 *Given a finite quantitative evidence frame $(S, \boldsymbol{\mathcal{E}}^Q)$, the min-max extension of $(\boldsymbol{\mathcal{E}}, \leq_e)$ is equivalent to $(2^{\boldsymbol{\mathcal{E}}} \setminus \{\emptyset\}, \preceq_e)$.*

Proof.

Let $\boldsymbol{E}$, $\boldsymbol{E}' \in 2^{\boldsymbol{\mathcal{E}}} \setminus \{\emptyset\}$. If for all $E \in \boldsymbol{E}$ there exists $E' \in \boldsymbol{E}'$ such that $E \leq_e E'$, then there exists $E' \in \boldsymbol{E}'$ such that $\max(\boldsymbol{E}) \leq_e E' \leq_e \max(\boldsymbol{E}')$. And, if for all $E' \in \boldsymbol{E}'$ there exists $E \in \boldsymbol{E}$ such that $E \leq_e E'$, then there exists $E \in \boldsymbol{E}$ such that $\min(\boldsymbol{E}) \leq_e E \leq_e \min(\boldsymbol{E}')$. So Definition 4.3 implies min-max extension. Conversely, if $\max(\boldsymbol{E}) \leq_e \max(\boldsymbol{E}')$ then $E \leq_e \max(\boldsymbol{E}')$ for every $E \in \boldsymbol{E}$. And if $\min(\boldsymbol{E}) \leq_e \min(\boldsymbol{E}')$ then $\min(\boldsymbol{E}) \leq_e E'$ for every $E' \in \boldsymbol{E}'$. □

Min-max extension is a pre-order satisfying the extension rule, extension dominance and extension independence, but not extension strict independence (Maly, 2020). For Proposition 4.3, we can extrapolate these properties to $(2^{\boldsymbol{\mathcal{E}}} \setminus \{\emptyset\}, \preceq_e)$.

Extension independence is a monotonicity rule, with respect to the certainty values, that would be counter-intuitive to violate in our context. If $\preceq_e$ would not satisfy extension independence then it could be the case that $\{E_1\} \preceq_e \{E_2\}$ but $\{E_2, E_3\} \preceq_e \{E_1, E_3\}$.

Something similar happens with violating the second statement of extension dominance. We can think of the case where $\{E_1\} \preceq_e \{E_2\}$ but $\{E_1, E_2\} \preceq_e \{E_1\}$. It is counter-intuitive to have more and more certain evidence in $\{E_1, E_2\}$ and still concluding that $\{E_1\}$ certainty-dominates it. The part of the principle that could be *a priori* rejected is Condition (1). However, rejecting Condition (1) implies that there will be cases where $\{E, E'\} \not\preceq_e \{E\}$ in spite of the fact that $q_{E'} < q_E$. Consequently, an order lifting of $(\boldsymbol{\mathcal{E}}, \leq_e)$ that does not satisfy dominance does not necessarily order the elements of $2^{\boldsymbol{\mathcal{E}}}$ according to the degrees of uncertainty of the individual evidence pieces in $\boldsymbol{\mathcal{E}}$: in this particular case, the resulting order $\preceq_e$ seems to ignore the fact that the only extra element E' of $\{E, E'\}$ is strictly less certain than E. Therefore, extension dominance and independence are indispensable properties for the comparison operator $\preceq_e$ for uncertain evidence.

This conclusion makes $(2^{\boldsymbol{\mathcal{E}}}, \preceq_e)$ specially relevant for the impossibility Theorems 4.1 and 4.2, and Proposition 4.4.

Theorem 4.1 (Kannai and Peleg, 1984). *Let X be a set with strictly more than five elements and $\leq$ a total order on X. Then there is no total pre-order on 2^X that satisfies extension dominance and extension independence with respect to $\leq$.*

Theorem 4.2 (Barberà and Pattanaik, 1984). *Let X be a set with strictly more than two elements and $\leq$ a total order on X. Then there is no binary relation on 2^X that satisfies extension dominance and extension strict independence with respect to $\leq$.*

Proposition 4.4 (Maly, 2020, Observation 3.1). *Let X be a set and $\leq$ a total order on X. Then every pre-order that satisfies extension dominance and extension independence includes the min-max extension.*

For Theorem 4.2, deciding to maintain extension dominance implies that we can only aim for extension (non-strict) independence. Since dropping extension independence is not an option in our context, Theorem 4.1 implies that any extension of $(\boldsymbol{\mathcal{E}}, \leq_e)$ we work with will be a non-total pre-order. Therefore, Proposition 4.4 determines that any other extension of $(\boldsymbol{\mathcal{E}}, \leq_e)$ that preserves extension dominance and extension independence will contain our order $(2^{\boldsymbol{\mathcal{E}}}, \preceq_e)$. This makes our proposal specially relevant for ordering sets of uncertain pieces of evidence. In Proposition 4.5 we collect all the order lifting properties of $(2^{\boldsymbol{\mathcal{E}}}, \preceq_e)$.

Proposition 4.5 *The following holds for any finite $(\boldsymbol{\mathcal{E}}, \leq_e)$ and the corresponding $(2^{\boldsymbol{\mathcal{E}}}, \preceq_e)$ (as given in Definition 4.3):*

1. *$(2^{\boldsymbol{\mathcal{E}}}, \preceq_e)$ satisfies the extension rule: for all $E, E' \in \boldsymbol{\mathcal{E}}$, if $E <_e E'$ then $\{E\} \prec_e \{E'\}$.*
2. *$(2^{\boldsymbol{\mathcal{E}}}, \preceq_e)$ satisfies extension dominance: for all $\boldsymbol{E} \subseteq \boldsymbol{\mathcal{E}}$ and all $E \in \boldsymbol{\mathcal{E}} \setminus \boldsymbol{E}$:*
   - *(a) if $E <_e E'$ for all $E' \in \boldsymbol{E}$, then $\boldsymbol{E} \cup \{E\} \prec_e \boldsymbol{E}$, and*
   - *(b) if $E' <_e E$ for all $E' \in \boldsymbol{E}$, then $\boldsymbol{E} \prec_e \boldsymbol{E} \cup \{E\}$.*
3. *$(2^{\boldsymbol{\mathcal{E}}}, \preceq_e)$ satisfies extension independence: for all $\boldsymbol{E}, \boldsymbol{E}' \subseteq \boldsymbol{\mathcal{E}}$ and $E \in \boldsymbol{\mathcal{E}} \setminus \boldsymbol{E} \cup \boldsymbol{E}'$, if $\boldsymbol{E} \prec_e \boldsymbol{E}'$, then $\boldsymbol{E} \cup \{E\} \preceq_e \boldsymbol{E}' \cup \{E\}$.*
4. *Every order extension $(2^{\boldsymbol{\mathcal{E}}}, \preceq)$ of $(\boldsymbol{\mathcal{E}}, \leq_e)$ such that (1) $\emptyset \preceq \boldsymbol{E}$ for every $\boldsymbol{E} \in 2^{\boldsymbol{\mathcal{E}}}$, (2) satisfies extension dominance and (3) extension independence, contains $(2^{\boldsymbol{\mathcal{E}}}, \preceq_e)$.*

Proof.

Since the restriction $(2^{\boldsymbol{\mathcal{E}}} \setminus \{\emptyset\}, \preceq_e)$ is exactly the min-max extension of $(\boldsymbol{\mathcal{E}}, \leq_e)$, all the statements hold for $\boldsymbol{E} \neq \emptyset$ (Maly, 2020). Including the empty set as the minimum of the order does not conflict with these statements, so they all hold for $(2^{\boldsymbol{\mathcal{E}}}, \preceq_e)$ as well. □

We propose $(2^{\boldsymbol{\mathcal{E}}}, \preceq_e)$ as a natural and relevant order extension of $(\boldsymbol{\mathcal{E}}, \leq_e)$ to capture uncertainty relations. Now, we will use $(2^{\boldsymbol{\mathcal{E}}}, \preceq_e)$ to compare propositions with respect to the degree of certainty of the evidential support they receive. Given a qualitative evidence frame $(S, \boldsymbol{\mathcal{E}}^Q)$ and $P \subseteq S$, we define the *basic evidence set* $\boldsymbol{\mathcal{E}}_P$ *for* P as $\boldsymbol{\mathcal{E}}_P = \{E \in \boldsymbol{\mathcal{E}} : E \subseteq P\}$. The order $\preceq_e$ is then straightforwardly extended to an order $\trianglelefteq_e$ on 2^S:

$$P \trianglelefteq_e Q \text{ if and only if } \boldsymbol{\mathcal{E}}_P \preceq_e \boldsymbol{\mathcal{E}}_Q. \qquad \text{(order } \trianglelefteq_e)$$

We will use this construction to compare propositions with respect to the degree of evidential support they receive, so let us present the properties of $(2^S, \trianglelefteq_e)$ in the following proposition.

Proposition 4.6 *Given a quantitative evidence frame* $(S, \boldsymbol{\mathcal{E}}^Q)$*, the pair* $(2^S, \trianglelefteq_e)$ *satisfies the following properties:*

(Reflexivity) $P \trianglelefteq_e P$ *for all* $P \in 2^S$.

(Transitivity) If $P_1 \trianglelefteq_e P_2$ *and* $P_2 \trianglelefteq_e P_3$ *then* $P_1 \trianglelefteq_e P_3$ *for all* $P_1, P_2, P_3 \in 2^S$.

(Minimum existence) $\emptyset$ *is the unique minimal element of* $(2^S, \trianglelefteq_e)$.

(Non-Monotonicity) $P \subseteq Q$ *does not imply* $P \trianglelefteq_e Q$.

(Extension rule) For all $E_1, E_2 \in \boldsymbol{\mathcal{E}}$ *and* $P, Q \subseteq S$ *such that* $\boldsymbol{\mathcal{E}}_P = \{E_1\}$ *and* $\boldsymbol{\mathcal{E}}_Q = \{E_2\}$*, if* $E_1 <_e E_2$ *then* $P \triangleleft_e Q$.

(Extension dominance) For all $P \subseteq S$ *and all* $E \in \boldsymbol{\mathcal{E}} \setminus \boldsymbol{\mathcal{E}}_P$*: (1) if* $E <_e E'$ *for all* $E' \in \boldsymbol{\mathcal{E}}_P$*, then for any* $Q \subseteq S$ *such that* $\boldsymbol{\mathcal{E}}_Q = \boldsymbol{\mathcal{E}}_P \cup \{E\}$, $Q \triangleleft_e P$*; (2) if* $E' <_e E$ *for all* $E' \in \boldsymbol{\mathcal{E}}_P$*, then for any* $Q \subseteq S$ *such that* $\boldsymbol{\mathcal{E}}_Q = \boldsymbol{\mathcal{E}}_P \cup \{E\}$, $P \triangleleft_e Q$.

(Extension independence) For all $P, Q \subseteq S$ *and* $E \in \boldsymbol{\mathcal{E}} \setminus (\boldsymbol{\mathcal{E}}_P \cup \boldsymbol{\mathcal{E}}_Q)$*, if* $P \triangleleft_e Q$ *then for any* $P', Q' \subseteq S$ *such that* $\boldsymbol{\mathcal{E}}_{P'} = \boldsymbol{\mathcal{E}}_P \cup \{E\}$ *and* $\boldsymbol{\mathcal{E}}_{Q'} = \boldsymbol{\mathcal{E}}_Q \cup \{E\}$, $P' \trianglelefteq_e Q'$

Proof.

All these properties are inherited from the order $(2^{\boldsymbol{\mathcal{E}}}, \preceq_e)$. □

4.3 The Logic

This section is dedicated to defining the logic for comparing strengths of evidence and belief. We introduce its syntax and semantics, discuss the nature of the modal operator for comparing uncertain evidence, and list some of its properties. Subsequently, we will propose a preliminary set of valid principles and point to important invalidities.

4.3.1 Syntax

We work with a binary modal language L with a countable set Prop of atomic formulas, $p, q, r, (p_1, p_2, \ldots)$, negation $\neg$, conjunction $\wedge$, a binary *evidence comparison modality* $\trianglelefteq_e$, a binary *belief comparison modality* $\trianglelefteq_b$ and round parentheses as auxiliary symbols. We use φ, ψ, χ $(\varphi_1, \varphi_2, ...)$, as metavariables for formulas of L. The well-formed formulas of our modal language L is given by the following grammar in BNF form:

$$A, B := p \mid \neg A \mid (A \wedge B)$$

$$\varphi, \psi := A \mid (A \trianglelefteq_e B) \mid (A \trianglelefteq_b B) \mid \neg\varphi \mid (\varphi \wedge \psi) \mid \Box\varphi$$

where $p \in \mathsf{Prop}$. We use $\rightarrow$ for the material conditional and $\vee$ for disjunction, defined in the usual manner as $\phi \vee \psi := \neg(\neg\phi \wedge \neg\psi)$ and $\phi \rightarrow \psi := \neg\phi \vee \psi$. We denote the propositional constants for tautology and contradiction by $\top$ and $\bot$, respectively. We will follow the usual rules for the elimination of the parentheses.

Notice that the first line of the BNF form defines the language of *classical propositional logic*, denoted henceforth by L_{CPL}. The comparison modalities $\trianglelefteq_e$ and $\trianglelefteq_b$ in L are intended to compare only first-order evidence and beliefs, thus, connect only the sentences in L_{CPL}. The intended reading of the binary modalities are as follows. We read $\varphi \trianglelefteq_e \psi$ as "*the evidence for ψ certainty-dominates the evidence for φ*;" and $\varphi \trianglelefteq_b \psi$ as "*belief in ψ is at least as strong as belief in φ*," following (Ghosh and De Jongh, 2013). The analogous belief comparison operator $\varphi \lhd \psi$ in (Harmanec and Hájek, 1994) is read as "*ψ is at least as believable as φ*" and interpreted with respect to a belief function in the same way we interpret $\trianglelefteq_b$. We here adopt the reading of the belief comparison operator $\succcurlyeq_B$ of (Ghosh and De Jongh, 2013) as we find it more intuitive and fitting to the formal interpretation and intended meaning of $\trianglelefteq_b$. $\Box\varphi$ is the epistemic modality "It is *a priori* that φ" and will be interpreted as the global modality.

We use the notation $\trianglelefteq_e$ both in syntax and semantics. In syntax it represents the evidence comparison modality. In semantics, we use it to *interpret* the evidence comparison modality. We believe it will be clear from the context when $\trianglelefteq_e$ is a syntactic object as a component of L and when it is part the semantics.

We introduce notations $\varphi \equiv_e \psi := (\varphi \trianglelefteq_e \psi) \wedge (\psi \trianglelefteq_e \varphi)$ and $\varphi \equiv_b \psi := (\varphi \trianglelefteq_b \psi) \wedge (\psi \trianglelefteq_b \varphi)$ for equivalence of φ and ψ with respects to evidential support and believability, respectively. We define the corresponding strong comparison operators $\triangleleft_e$ and $\triangleleft_b$, respectively as, $\varphi \triangleleft_e \psi := (\varphi \trianglelefteq_e \psi) \wedge \neg(\psi \trianglelefteq_e \varphi)$ and $\varphi \triangleleft_b \psi := (\varphi \trianglelefteq_b \psi) \wedge \neg(\psi \trianglelefteq_b \varphi)$. We read $\varphi \triangleleft_e \psi$ as "*the evidence for ψ strictly certainty-dominates the evidence for φ*;" and $\varphi \triangleleft_b \psi$ as "*belief in ψ is strictly stronger than belief in φ*".

4.3.2 Semantics

In this subsection we define the semantics of our logic and compare our models with some of those proposed in relevant literature.

Definition 4.6 (Quantitative Evidence-Belief Models). A *quantitative evidence-belief model* (or, in short, an *e-b model*) $\mathcal{M} = \langle S, \boldsymbol{\mathcal{E}}^Q, \mathrm{Bel}, V\rangle$ is a tuple, where

1. $(S, \boldsymbol{\mathcal{E}}^Q)$ is quantitative evidence frame,
2. $\mathrm{Bel} : 2^S \to [0, 1]$ is a belief function computed by the multi-layer belief model with input $\boldsymbol{\mathcal{E}}^Q$, Dempster-Shafer frame of justification and the intersection as allocation function. In other words, Bel is the belief function obtained from applying the Dempster's rule of combination to the collection of basic belief assignments m_j such that $E_j \in \boldsymbol{\mathcal{E}}$ is the only proper focal element of m_j and $m_j(E_j) = q_j$ for every $j \in \{1, \dots, \ell\}$. Recall that Bel is $\mathrm{Bel}_{\mathcal{J}^{DS}}(i, \cdot)$ from Chapter 3.
3. $V : \mathsf{Prop} \to 2^S$ is a standardly defined valuation map.

Our e-b models are a combination of the topological models of evidence of (Baltag et al., 2022) and qualitative belief models of (Harmanec and Hájek, 1994). Topological models of evidence are defined as $\mathcal{M} = \langle S, \boldsymbol{\mathcal{E}}, \tau, V\rangle$, where S is a set of possible states, $\boldsymbol{\mathcal{E}} \subseteq 2^S$, τ is the tautology generated by $\boldsymbol{\mathcal{E}}$ and V a valuation map. Similarly to e-b models, these models provide an explicit representation of evidence, as the set $\boldsymbol{\mathcal{E}}$ represents a collection of partial and possibly mutually contradictory evidence. In addition, both settings state $\emptyset \notin \boldsymbol{\mathcal{E}}$. One of the main differences is that $\boldsymbol{\mathcal{E}}$ in topological models of evidence always contains the total set S, while in our context $S \notin \boldsymbol{\mathcal{E}}$ for its role representing uncertainty. More substantially, topological models of evidence are purely qualitative, lacking quantitative components such as Q and Bel. This means that topological models of evidence do not represent uncertainty about evidence. As opposed to the topological evidence models, e-b models do not have an explicit representation of evidential topology, that is, an explicit representation of a set of combined evidence pieces and arguments (although these can be generated from $\boldsymbol{\mathcal{E}}$).

Qualitative belief models of (Harmanec and Hájek, 1994), on the other hand, are defined as $\mathcal{M} = \langle S, \mathrm{Bel}, V\rangle$, where S is a set of possible states, Bel is a belief function and V a valuation function. They are similar to e-b models in that they both have a representation of degrees of belief based on belief functions. The main difference is that qualitative belief models do not include an explicit representation of evidence, and the function Bel is not necessarily computed through Dempster's rule of combination or any other specific combination rule. In qualitative belief models, Bel is a generic belief function, whereas in e-b models, it is a separable belief function. Our e-b models put together the topological models of evidence of (Baltag et al., 2022) and qualitative belief models of (Harmanec and Hájek, 1994) by enriching the former via components that represent degrees of uncertainty about evidence and belief, and the latter by an explicit representation of pieces of evidence.

To interpret the modality $\trianglelefteq_e$ in the intended way, we use the pre-order that we denoted by the same symbol $\trianglelefteq_e$ in Equation order $\trianglelefteq_e$. The semantics $\models$ for L in e-b models is defined recursively as in Definition 4.7. The *truth set* of $\varphi \in$ L (or, the *intension* of φ) with respect to $\mathcal{M}$ is $[\![\varphi]\!]_{\mathcal{M}} := \{s \in S : \mathcal{M}, s \models \varphi\}$, namely, the set of all possible states that makes φ true. We omit the subscript $\mathcal{M}$ in $[\![\varphi]\!]_{\mathcal{M}}$ when the model is contextually clear. To simplify notation, instead of writing $\boldsymbol{\mathcal{E}}_{[\![\varphi]\!]_{\mathcal{M}}}$, we simply write $\boldsymbol{\mathcal{E}}_{\varphi}$ when the model is clear from context and call $\boldsymbol{\mathcal{E}}_{\varphi}$ the *evidence set for* φ.

Definition 4.7 (Semantics for L ($\models$)). Given an e-b model $\mathcal{M} = \langle S, \boldsymbol{\mathcal{E}}^Q, \mathrm{Bel}, V\rangle$ and a state $s \in S$, the $\models$-semantics for the language L is defined recursively as:

$$
\begin{array}{lll}
\mathcal{M}, s \models p & \text{if and only if} & s \in V(p) \\
\mathcal{M}, s \models \neg\varphi & \text{if and only if} & \text{not } \mathcal{M}, s \models \varphi \\
\mathcal{M}, s \models \varphi \wedge \psi & \text{if and only if} & \mathcal{M}, s \models \varphi \text{ and } \mathcal{M}, s \models \psi \\
\mathcal{M}, s \models \varphi \trianglelefteq_e \psi & \text{if and only if} & \boldsymbol{\mathcal{E}}_{\varphi} \trianglelefteq_e \boldsymbol{\mathcal{E}}_{\psi} \\
\mathcal{M}, s \models \varphi \trianglelefteq_b \psi & \text{if and only if} & \mathrm{Bel}([\![\varphi]\!]) \leq \mathrm{Bel}([\![\psi]\!]) \\
\mathcal{M}, s \models \Box\varphi & \text{if and only if} & S \subseteq [\![\psi]\!].
\end{array}
$$

where $p \in \mathsf{Prop}$.

The notions of logical consequence and validity are defined standardly as follows. Given a $\Gamma \subseteq$ L and $\varphi \in$ L, we say that φ is a *logical consequence* of Γ, denoted by $\Gamma \models \varphi$, if for all e-b models $\mathcal{M} = \langle S, \boldsymbol{\mathcal{E}}^Q, \mathrm{Bel}, V\rangle$ and all $s \in S$: if $\mathcal{M}, s \models \psi$ for all $\psi \in \Gamma$, then $\mathcal{M}, s \models \varphi$. For single-premise entailment, we write $\psi \models \varphi$ for $\{\psi\} \models \varphi$. As a special case, *validity*, $\models \varphi$, is truth at all states of all e-b models (equivalently, it is entailment by the empty set of premises). φ is called invalid, denoted by $\not\models \varphi$, if it is not a validity, that is, if there is an e-b model $\mathcal{M} = \langle S, \boldsymbol{\mathcal{E}}^Q, \mathrm{Bel}, V\rangle$ and a state $s \in S$ such that $\mathcal{M}, s \not\models \varphi$ (Blackburn et al., 2001).

While the semantics for the Booleans are standard, the semantic clause for the belief comparison operator $\trianglelefteq_b$ is the same as the one proposed in (Harmanec and Hájek, 1994) (operator $\lhd$ in their notation). In light of Proposition 3.8, the logic of the subfragment of L without $\trianglelefteq_e$ (equivalently, having only $\trianglelefteq_b$ as a modal operator) includes the logic presented in (Harmanec and Hájek, 1994) since the belief function Bel in an e-b model $\mathcal{M} = \langle S, \boldsymbol{\mathcal{E}}^Q, \mathrm{Bel}, V\rangle$ is not any belief function as in the models of (Harmanec and Hájek, 1994), but it is, in particular, a separable support function.

The novel bit of the logic is the evidence comparison operator $\trianglelefteq_e$ and, therefore, it deserves further elaboration. According to our semantics, *evidence for ψ certainty-dominates the evidence for φ* if and only if the evidence set for ψ is *at least as certain as* the evidence set of φ, that is, either there is no evidence for φ or (1) every piece of evidence for φ is certainty-dominated by some piece of evidence for ψ, and (2) every piece of evidence for ψ certainty-dominates some piece of evidence for φ.

In Section 4.2, we saw that this definition preserves the notion of certainty from pieces of evidence in $\boldsymbol{\mathcal{E}}$ to propositions in 2^S through the extension dominance property (Proposition 4.6). In addition, any comparison operator defined as an order lifting of $(\boldsymbol{\mathcal{E}}, \leq_e)$ and satisfying extension dominance will include the relations established by $\trianglelefteq_e$. Therefore, the operator $\trianglelefteq_e$ is a good comparison operator for uncertain evidence, regardless of the combination method we use to compute Bel.

The next part of the section is dedicated to exploring the behavior of this operator and its belief counterpart $\trianglelefteq_b$. We also point to a few connections between the two operators.

4.3.3 (In)Validities and Expressivity

In this section we provide a list of interesting validities of the proposed logic, have a look at some important properties about evidence, belief, and their connection that the language L can express. Finally, we also point to some common and distinguishing features of the proposed framework and its close relatives such as the logics of (Harmanec and Hájek, 1994), (Ghosh and De Jongh, 2013), and (Baltag et al., 2022).

Properties of Belief Order ($\trianglelefteq_b$)

Proposition 4.7 *The following principles are valid in all e-b models:*

B1 $\varphi \trianglelefteq_b \varphi$

B2 $(\varphi \trianglelefteq_b \psi) \vee (\psi \trianglelefteq_b \varphi)$

B3 $((\varphi \trianglelefteq_b \psi) \wedge (\psi \trianglelefteq_b \chi)) \to (\varphi \trianglelefteq_b \chi)$

B4 $\neg(\top \trianglelefteq_b \bot)$

B5 $\Box(\varphi \to \psi) \to (\varphi \trianglelefteq_b \psi)$

B6 $(\Box(\varphi \to \psi) \wedge \Box\neg(\psi \wedge \chi) \wedge \neg(\psi \trianglelefteq_b \varphi)) \to \neg((\psi \vee \chi) \trianglelefteq_b (\varphi \vee \chi))$

B7 $(\Box\varphi \wedge \neg\Box\psi) \to \neg(\varphi \trianglelefteq_b \psi)$

B8 $(\varphi \trianglelefteq_b \psi) \to \Box(\varphi \trianglelefteq_b \psi)$

B9 $\neg(\varphi \trianglelefteq_b \psi) \to \Box\neg(\varphi \trianglelefteq_b \psi)$

Proof.

B1 follows from the fact that $([0,1], \leq)$ is a total order, thus, induces a reflexive order on 2^S.

B2 follows from the fact that $([0,1], \leq)$ is a total order, thus, induces a total order on 2^S.

B3 follows from the fact that $([0,1], \leq)$ is a total order, thus, induces a transitive order on 2^S.

B4 follows from the properties of Bel such that $\mathrm{Bel}(\llbracket\top\rrbracket) = 1 \not\leq \mathrm{Bel}(\llbracket\bot\rrbracket) = 0$.

To prove B5, suppose $\mathcal{M}, s \models \Box(\varphi \to \psi)$. This means that $\llbracket\varphi\rrbracket \subseteq \llbracket\psi\rrbracket$. Therefore, for all $E \in \tau_{\mathcal{E}}$ such that $E \subseteq \llbracket\varphi\rrbracket$, we have $E \subseteq \llbracket\psi\rrbracket$. Therefore, $\mathrm{Bel}(\llbracket\varphi\rrbracket) \leq \mathrm{Bel}(\llbracket\psi\rrbracket)$.

B6 follows from the axiom of partial monotonicity ($A \subseteq B$, $B \cap C = \emptyset$ and $\mathrm{Bel}(B) > \mathrm{Bel}(A)$ implies $Bel(B \cup C) > Bel(A \cup C)$) for all belief functions Bel used in (Wong et al., 1991) to characterize a belief function that fully agrees with a preference relation. If $\mathcal{M}, s \models (\Box(\varphi \to \psi) \wedge \Box\neg(\psi \wedge \chi) \wedge \neg(\psi \trianglelefteq_b \varphi))$, then $\llbracket\varphi\rrbracket \subseteq \llbracket\psi\rrbracket$, $\llbracket\psi\rrbracket \cap \llbracket\chi\rrbracket = \emptyset$ and $\mathrm{Bel}(\llbracket\varphi\rrbracket) < \mathrm{Bel}(\llbracket\psi\rrbracket)$. By the partial monotonicity axiom, this implies that $\mathrm{Bel}(\llbracket\varphi\rrbracket \cup \llbracket\chi\rrbracket) < \mathrm{Bel}(\llbracket\psi\rrbracket \cup \llbracket\chi\rrbracket)$. Since $\llbracket\varphi \vee \chi\rrbracket = \llbracket\varphi\rrbracket \cup \llbracket\chi\rrbracket$ and $\llbracket\psi \vee \chi\rrbracket = \llbracket\psi\rrbracket \cup \llbracket\chi\rrbracket$, we conclude $\mathcal{M}, s \models \neg((\psi \vee \chi) \trianglelefteq_b (\varphi \vee \chi))$.

To prove B7, suppose $\mathcal{M}, s \models \Box\varphi \wedge \neg\Box\psi$. This means that $S \subseteq \llbracket\varphi\rrbracket$ and $S \not\subseteq \llbracket\psi\rrbracket$. The former entails that $S = \llbracket\varphi\rrbracket$. As $\mathrm{Bel}(P) \in [0,1]$ for all $P \subseteq S$ and $\mathrm{Bel}(S) = 1$, we obtain that $\mathrm{Bel}(\llbracket\psi\rrbracket) \leq \mathrm{Bel}(\llbracket\varphi\rrbracket) = 1$. For our assumption of $q_E \neq 1$ for every $E \in \mathcal{E}$, $\mathrm{Bel}(P) < 1$ for every P such that $\llbracket P\rrbracket \subsetneq S$. Therefore, $\mathrm{Bel}(\llbracket\psi\rrbracket) < \mathrm{Bel}(\llbracket\varphi\rrbracket) = 1$ and $\mathcal{M}, s \models \neg(\varphi \trianglelefteq_b \psi)$.

Formulas B8 and B9 simply say that belief order is state independent.

□

These formulas describe the properties of $\trianglelefteq_b$ as a total order defined according to a belief function. Validities B2, B3,B4, B5, B6, B8 and B9 define the logic in (Harmanec and Hájek, 1994) (although we interpret □ differently). In addition, validities B1, B2, B3 and B5 are the axioms from the logic for comparing strength of belief $KD45-O$ of (Ghosh and De Jongh, 2013) that only include comparison operator analogous to $\trianglelefteq_b$. Furthermore, validity B7 is inspired by the logic for safe belief $KD45-OS$ defined in (Ghosh and De Jongh, 2013) (where their $\Box$ operator is an $S4$ modality, rather than the global modal operator as in our case). In our case, this formula is true only under our restriction on not accepting dogmatic evidence, that is, our constraint on $q \in Q$ being strictly lower than 1. Similarly, validity B4 encodes non-triviliaty of $\boldsymbol{\mathcal{E}}$—i.e., $\boldsymbol{\mathcal{E}} \neq \emptyset$—and validities B5 and B6 refers to the additive nature of Bel—as $\mathrm{Bel}_{\mathcal{J}^{DS}}(i, P)$ is the addition of the masses $\delta_{\mathcal{J}^{DS}}(i, P')$ of the subsets P' of P.

Properties of Evidence Order ($\trianglelefteq_e$)

Proposition 4.8 *The following principles are valid in all e-b models:*

E1 $\varphi \trianglelefteq_e \varphi$.

E2 $((\varphi \trianglelefteq_e \psi) \wedge (\psi \trianglelefteq_e \chi)) \rightarrow (\varphi \trianglelefteq_e \chi)$.

E3 $\bot \trianglelefteq_e \varphi$

E4 $\neg(\top \trianglelefteq_e \bot)$.

E5 $\Box(\varphi \leftrightarrow \psi) \rightarrow (\varphi \equiv_e \psi)$

E6 $((\varphi \trianglelefteq_e \chi) \wedge (\psi \trianglelefteq_e \chi) \wedge \Box(\varphi \rightarrow \eta) \wedge \Box(\eta \rightarrow \psi)) \rightarrow (\eta \trianglelefteq_e \chi)$

E7 $((\chi \trianglelefteq_e \varphi) \wedge (\chi \trianglelefteq_e \psi) \wedge \Box(\varphi \rightarrow \eta) \wedge \Box(\eta \rightarrow \psi)) \rightarrow (\chi \trianglelefteq_e \eta)$

Proof.

E1 follows from the fact that $\trianglelefteq_e$ is a pre-order on $2^{\boldsymbol{\mathcal{E}}}$, thus, it is in particular reflexive (Proposition 4.6).

E2 follows from the fact that $\trianglelefteq_e$ is a pre-order on $2^{\boldsymbol{\mathcal{E}}}$, thus, it is in particular transitive (Proposition 4.6).

E3 follows from the fact that $\emptyset \trianglelefteq_e \boldsymbol{E}$ for all $E \in 2^{\boldsymbol{\mathcal{E}}}$.

E4 follows from the fact that $\boldsymbol{\mathcal{E}} \neq \emptyset$.

To prove E5, suppose $\mathcal{M}, s \models \Box(\varphi \leftrightarrow \psi)$. This means that $[\![\varphi]\!] = [\![\psi]\!]$. The latter means that $\boldsymbol{\mathcal{E}}_\varphi = \boldsymbol{\mathcal{E}}_\psi$, thus, $\mathcal{M}, s \models (\varphi \equiv_e \psi)$.

To prove E6, suppose that $\mathcal{M}, s \models \Box(\varphi \rightarrow \eta) \wedge \Box(\eta \rightarrow \psi)$. This means that $[\![\varphi]\!] \subseteq [\![\eta]\!] \subseteq [\![\psi]\!]$. Therefore, $\min(\{q_E : E \in [\![\psi]\!]\}) \leq \min(\{q_E : E \in [\![\eta]\!]\}) \leq \min(\{q_E : E \in [\![\varphi]\!]\})$ and $\max(\{q_E : E \in [\![\varphi]\!]\}) \leq \max(\{q_E : E \in [\![\eta]\!]\}) \leq \max(\{q_E : E \in [\![\psi]\!]\})$. Now, suppose that $\mathcal{M}, s \models (\varphi \trianglelefteq_e \chi) \wedge (\psi \trianglelefteq_e \chi)$. This means, following the min-max definition of $\trianglelefteq_e$ (Definition 4.5), that $\min(\{q_E : E \in [\![\varphi]\!]\}) \leq \min(\{q_E : E \in [\![\chi]\!]\})$ and $\max(\{q_E : E \in [\![\psi]\!]\}) \leq$

$\max(\{q_E : E \in [\![\chi]\!]\})$. Merging these two facts, we get that $\min(\{q_E : E \in [\![\eta]\!]\}) \leq \min(\{q_E : E \in [\![\chi]\!]\})$, $\max(\{q_E : E \in [\![\eta]\!]\}) \leq \max(\{q_E : E \in [\![\chi]\!]\})$, thus, $\mathcal{M}, s \models \eta \trianglelefteq_e \chi$.

E7 follows from a similar reasoning as before, noticing that $\mathcal{M}, s \models (\chi \trianglelefteq_e \varphi) \wedge (\chi \trianglelefteq_e \psi)$ implies that $\min(\{q_E : E \in [\![\chi]\!]\}) \leq \min(\{q_E : E \in [\![\psi]\!]\})$ and $\max(\{q_E : E \in [\![\chi]\!]\}) \leq \max(\{q_E : E \in [\![\varphi]\!]\})$. $\square$

Validities E1 and E2 capture the order properties reflexivity and transitivity of $(2^S, \trianglelefteq_e)$. Validities E3 and E4 are specific for evidence models, since they state that the empty set is never evidentially supported and non-triviality of the evidence set respectively. Validities E5, E6 and E7 come from the logic of convex order in (Shi and Sun, 2021, p.1022). Validity E5 states that necessarily equivalent propositions are supported by the evidence to equal degree. Notice that the stronger principle $\Box(\varphi \to \psi) \to (\varphi \trianglelefteq_e \psi)$ is not valid. This is because $(2^S, \trianglelefteq_e)$ is a non-monotonic order (see Proposition 4.6 and proof of Proposition 4.2). Finally, validities E6 and E7 bring some uniformity to the uncertainty comparison within nested sets. These validities show that the smallest and biggest sets of a nested chain of sets determine the upper and lower bounds for the certainty-dominance of the intermediate sets.

Before we list the connecting principles, we state some of the important properties our language can express:

$\neg(\varphi \trianglelefteq_b \bot)$: φ is believed to some non-zero degree.

$$\begin{aligned}
\mathcal{M}, s \models \neg(\varphi \trianglelefteq_b \bot) \quad & \text{if and only if} \quad \text{not } (\mathrm{Bel}([\![\varphi]\!]) \leq \mathrm{Bel}([\![\bot]\!])) \\
& \text{if and only if} \quad \text{not } (\mathrm{Bel}([\![\varphi]\!]) \leq 0) \qquad ([\![\bot]\!] = \emptyset, \mathrm{Bel}(\emptyset) = 0) \\
& \text{if and only if} \quad 0 < \mathrm{Bel}([\![\varphi]\!]) \qquad (\leq \text{ real order on } [0,1])
\end{aligned}$$

$\neg(\varphi \trianglelefteq_e \bot)$: there is a basic piece of evidence for φ.

$$\begin{aligned}
\mathcal{M}, s \models \neg(\varphi \trianglelefteq_e \bot) \quad & \text{if and only if} \quad \text{not } (\boldsymbol{\mathcal{E}}_\varphi \trianglelefteq_e \boldsymbol{\mathcal{E}}_\bot) \\
& \text{if and only if} \quad \text{not } (\boldsymbol{\mathcal{E}}_\varphi \trianglelefteq_e \emptyset) \qquad (\boldsymbol{\mathcal{E}}_\bot = \emptyset) \\
& \text{if and only if} \quad \boldsymbol{\mathcal{E}}_\varphi \neq \emptyset \qquad (\text{Prop. 4.2.2 and Def.4.3.1})
\end{aligned}$$

Last formula is the modality E_0 in (Baltag et al., 2022).

Connecting Principles

Proposition 4.9 *The following principle is valid in all e-b models:*

C1 $\neg(\varphi \trianglelefteq_e \bot) \to \neg(\varphi \trianglelefteq_b \bot)$

Proof.

C1 follows from the fact that $0 < q_E < 1$ for every $E \in \boldsymbol{\mathcal{E}}$ implies that $\delta_{\mathcal{J}^{DS}}(i, E) > 0$ for every $E \in \boldsymbol{\mathcal{E}}$ and therefore $\mathrm{Bel}_{\mathcal{J}^{DS}}(i, E) > 0$. □

Validity C1 means that whatever is supported by some evidence will be believed to some non-zero degree by the agent. This is the case only for our restriction to non-dogmatic evidence. If $A \in \boldsymbol{\mathcal{E}}$ and $q_A = 1$, then $\mathrm{Bel}_{\mathcal{J}^{DS}}(i, B) = 0$ for every $B \cap A = \emptyset$. This connecting principle relates basic evidence and belief. However, it does not establish a connection that reveals the combination rule that links $\boldsymbol{\mathcal{E}}^Q$ and Bel in the model. The restriction of Bel to support functions is caught in some invalidities of the models instead.

Important invalidities **Proposition 4.10** *The following principles are invalid in e-b models:*

I1 $\not\models \neg(\varphi \trianglelefteq_b \bot) \to \neg(\varphi \trianglelefteq_e \bot)$

I2 $\not\models \Box(\varphi \to \psi) \to (\varphi \trianglelefteq_e \psi)$

I3 $\not\models \neg(\varphi \trianglelefteq_b \psi) \to \neg(\neg\psi \trianglelefteq_b \neg\varphi)$

Proof.

To prove I1, consider the e-b model $\mathcal{M} = \langle S, \boldsymbol{\mathcal{E}}^Q, \mathrm{Bel}, V\rangle$ such that $S = \{a, b, c\}$, $\boldsymbol{\mathcal{E}}^Q = \Big\{(\{a, b\}, 0.7), (\{b, c\}, 0.6)\Big\}$, and $V(p) = \{b\}$. It holds that $\mathrm{Bel}([\![p]\!]) > 0$ but there is no $E \in \boldsymbol{\mathcal{E}}$ such that $E \subseteq \{b\}$. Therefore, $\mathcal{M}, a \not\models \neg(p \trianglelefteq_b \bot) \to \neg(p \trianglelefteq_e \bot)$.

To prove I2, consider the e-b model $\mathcal{M} = \langle S, \boldsymbol{\mathcal{E}}^Q, \mathrm{Bel}, V\rangle$ such that such that $S = \{a, b, c\}$ and $\boldsymbol{\mathcal{E}}^Q = \Big\{(\{a\}, 0.7), (\{a, b\}, 0.4)\Big\}$, $V(p) = \{a\}$ and $V(q) = \{a, b\}$. Then, obviously $\mathcal{M}, a \models \Box(p \to q)$. However, $\mathcal{M}, a \not\models p \trianglelefteq_e q$ since $\{a, b\} \in \boldsymbol{\mathcal{E}}_q$ and there is no $E \in \boldsymbol{\mathcal{E}}_p$ such that $E \leq_e \{a, b\}$.

To show I3, consider the previous model again. $\mathcal{M}, a \models \neg(p \trianglelefteq_b q)$, since $\mathrm{Bel}(\{a, b\}) = \delta_{\mathcal{J}^{DS}}(i, \{a, b\}) + \delta_{\mathcal{J}^{DS}}(i, \{a\}) > \delta_{\mathcal{J}^{DS}}(i, \{a\}) = \mathrm{Bel}(\{a\})$. However, $\mathcal{M}, a \not\models \neg(\neg q \trianglelefteq_b \neg p)$, because $\mathrm{Bel}(\{c\}) = \mathrm{Bel}(\{b, c\}) = 0$. □

Invalidity I1 means that believing in a proposition does not guarantee that the agent possesses basic evidence supporting it. This belief may be supported by the combined evidence obtained after running the multi-layer belief model. Therefore, we cannot infer the existence of basic evidence from a strictly stronger belief. Invalidity I2 comes from the non-monotonic nature of $\trianglelefteq_e$, which allows us to differenciate certainty degrees beyond entailment (see discussion in Section 4.2). Lastly, invalidity I3 remarks the non-probabilistic character of belief functions.

This formula is an optional axiom for the logic $KD45 - O$ of (Ghosh and De Jongh, 2013). In our case this axiom is not valid because Bel may give values to P and $S \setminus P$ such that $\mathrm{Bel}(P) + \mathrm{Bel}(S \setminus P) \neq 1$.

Deduction rules The following inference rules are validity preserving:

1. (MP) From $\varphi \to \psi$, φ infer ψ.
2. Necessity of $\Box$: From φ infer derives $\Box\varphi$.

Comparison with reference logics Throughout this section, we have seen connections and disagreements with the logics of reference presented in (Harmanec and Hájek, 1994), (Shi and Sun, 2021) and (Ghosh and De Jongh, 2013). Now, we conclude with a detailed discussion of the similarities and dissimilarities of these logics with ours.

First, let us focus on the qualitative belief logic (QBL) of (Harmanec and Hájek, 1994). This logic shares the explicit comparison operator for belief, but it is not expressive enough to reason about the evidence that support that belief. As seen above, QBL is fully contained in ours, since it is defined by the validities B2, B3,B4, B5, B6, B8 and B9. Our main contributions with respect to QBL is that the function Bel of QBL models is a belief function in general and not necessarily a support function. Therefore, if this logic had a representation for evidence, it would represent the basic belief assignments corresponding to Bel, but not a collection of basic belief assignments to be combined. In that case, the formula I1 would be valid with respect to evidence models of QBL, whereas it is not valid in ours. In any case, a next step would be to strength our logic to be able to capture the difference between basic and combined evidence in some validities and eliminate our restriction to separable support functions.

Another reference logic for us is the logic of convex order (AC) defined in (Shi and Sun, 2021). This logic has a comparison operator defined in terms of Egli-Milner order, which links it to our operator $\trianglelefteq_e$. Their comparison operator is neither connected to the notion of evidence nor linked to other modal operators that may refer to belief. Despite the direct influence of Egli-Milner order in both operators, our logic and the axiom system AC for logic of convex order differ notably. One difference is that we impose the empty set as the minimum of the order $(2^S, \trianglelefteq_e)$ (Definitions 4.3 and order $\trianglelefteq_e$). This makes axiom BT of the axiom system AC—i.e., the axiom $(\bot \trianglelefteq_e \varphi) \to (\varphi \trianglelefteq_e \bot)$—invalid in our logic. For validity E3, we have $\bot \trianglelefteq_e \varphi$ as a validity, but $\varphi \trianglelefteq_e \bot$ is not satisfied in those models $\mathcal{M}$ such that $[\![\varphi]\!]_{\mathcal{M}} \neq \emptyset$. Another difference is that Shi and Sun (2021) extend a preference order on S to 2^S directly. However, we move from an order on sets of possible states $(\boldsymbol{\mathcal{E}}, \leq_e)$ to an order of sets of these sets $(2^{\boldsymbol{\mathcal{E}}}, \preceq_e)$. From this order, we easily define an order on 2^S based on the notion of evidential support for propositions, but this intermediate step introduce new behaviours in the resulting logic. In particular, the axiom $J+$ of the axiom system AC—i.e., $((\varphi \trianglelefteq_e \psi) \wedge (\chi \trianglelefteq_e \eta)) \rightarrow ((\varphi \vee \chi) \trianglelefteq_e (\psi \vee \eta))$—is not valid in our logic. For example, let $\mathcal{M} = \langle S, \boldsymbol{\mathcal{E}}^Q, \mathrm{Bel}, V\rangle$ be a e-b- model such that $S = \{a, b, c, d\}$ and $\boldsymbol{\mathcal{E}}^Q = \Big\{(\{a, c\}, 0.6), (\{b\}, 0.3), (\{d\}, 0.2)\Big\}$, and $V(p) = \{a\}$, $V(q) = \{b\}$, $V(r) = \{c\}$ and $V(s) = \{d\}$, we get that $\mathcal{M} \models (p \trianglelefteq_e q) \wedge (r \trianglelefteq_e s)$ but $\mathcal{M} \not\models (p \vee r) \trianglelefteq_e (q \vee s)$. The other four axioms of the system AC are the validities E5, E2, E6 and E7.

The last logic we took as a reference is the $KD45 - O$ logic defined in (Ghosh and De Jongh, 2013). This logic defines a comparison operator for belief that is read the same way as our $\trianglelefteq_b$ (and intended to capture a similar notions of strength of belief) and a unary modal operator B for belief. In our case, we could introduce a unary belief operator as $\neg(\varphi \trianglelefteq_b \bot)$ or by defining belief above certain threshold. In addition, they also have the global modality in their logic ($\Box$ in our notation, and U in theirs). The logic $KD45 - O$ is defined by nine axioms. Two of them establish introspection relations that we are not able to express in our language. The other five, are all valid in our logic. They correspond to validities B1, B2, B3, B5 and $(\neg(\varphi \trianglelefteq_b \bot) \wedge (\psi \trianglelefteq_b \bot)) \rightarrow \neg(\varphi \trianglelefteq_b \psi)$, which follows from B2 and B3. This means that, $KD45 - O$ is formally not contained in our logic, but the common and relevant part is contained both in our logic and QBL.

We conclude our analysis here, hoping that it serves as a first step in developing a logic for comparing the strengths of evidence and belief based on e-b models. While the presented logic contributes to the literature by establishing an explicit link between uncertain evidence and its corresponding multi-layer belief functions, some further work is needed to strengthen these links and to identify validities that may change with respect to the parameterization of the multi-layer belief model. Furthermore, the logic of e-b models can be enriched by introducing the notion of justification or by making the language expressible enough to identify combined evidence. We conclude this chapter with a use case scenario that highlights the importance of reasoning about the different components involved in computing belief functions, such as the degree of certainty of the evidence pieces, and thus motivate this chapter from a different perspective.

4.4 Use Case Scenario: Science Communication

Explicitly reasoning about evidence certainty and the degree of belief that this evidence generates has a clear theoretical benefit. A logic like the one presented in this chapter allows us to identify mathematical properties to describe evidence combination methods beyond a specific scenario. However, pursuing this research line can also be supported by real-life motivations.

The multi-layer belief model aggregates evidence according to its certainty values, the number of pieces of evidence—since the degree of belief is defined as an addition—and certain assumptions about justifications and evidence allocation functions. Therefore, computing the final degree of belief takes into account various relevant features of evidence, but can be also seen as a black box from the final user's point of view. In other words, we know that uncertainty, the number of pieces of evidence, and consistency of the evidence are affecting the degree of belief the multi-layer belief model assigns to propositions, but we do not know which feature, if any, is affecting the result to what extent.

To illustrate how our logic may clarify some practical problems, let us consider a situation where summarizing measures may streamline decision making, but transparency is equally, if not more, important. The COVID-19 pandemic of 2020 showed that science-driven policies, together with science communication, may represent one of these scenarios. During the pandemic, public health institutions had to make many exceptional decisions with a huge impact on the population, based on partial and uncertain evidence. Furthermore, the effectiveness of those decisions directly depended on public engagement, making not only the recommendation itself but also its justification critically important.

Imagine it is December 2020, and the World Health Organization must make a recommendation about vaccination for COVID-19. Let us say that the organization counts on the following quantitative evidence set:

$$(\{\text{No Symptoms}\}, 0.9)$$
$$(\{\text{No Symptoms}, \text{Minor Symptoms}\}, 0.7)$$
$$(\{\text{Severe Symptoms}\}, 0.2)$$

where $\{\text{No Symptoms}\}$, $\{\text{No Symptoms}, \text{Minor Symptoms}\}$, and $\{\text{Severe Symptoms}\}$ respectively refer to evidence for 'vaccination against COVID-19 prevents all symptoms', 'vaccination against COVID-19 prevents severe symptoms', and 'vaccination against COVID-19 causes side effects with severe symptoms'. For simplicity, we denote No Symptoms by NS, Minor Symptoms by MS, and Severe Symptoms by SS. So let $S = \{NS, MS, SS\}$ be a set of possible states, and $\boldsymbol{\mathcal{E}}^Q = \Big\{(\{NS\}, 0.9), (\{NS, MS\}, 0.7), (\{SS\}, 0.2)\Big\}$ be a quantitative evidence set.

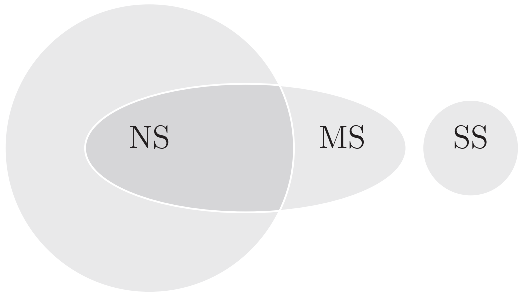


Figure 4.5: Representation of uncertain pieces of evidence in the running example by a Venn diagram. The size of the area represents the degree of certainty of the corresponding piece of evidence.

In this scenario, the World Health Organization could aggregate their evidence by applying Dempster's rule of combination, obtaining $\mathrm{Bel}(\{NS, MS\}) = 0.96$ and $\mathrm{Bel}(\{SS\}) = 0.01$, and conclude that worldwide vaccination should be recommended. However, regardless of the theoretical soundness of this combination method for evidence, stating that 'this recommendation is supported by a 0.96 degree of belief' is meaningless for the general population. Questions such as *what degree of belief means* or *how it is computed* may trigger distrust among citizens.

Alternatively, assume that the World Health Organization has a model checker for the logic introduced in this chapter. Let us imagine that this model checker states that, for our particular S and $\boldsymbol{\mathcal{E}}^Q$, the model $\mathcal{M} = \langle S, \boldsymbol{\mathcal{E}}^Q, \mathrm{Bel}, V\rangle$ satisfies formula (4.1)—which reads as *if* ψ *certainty-dominates* φ *then belief in* ψ *is at least as strong as belief in* φ—in every possible world. Then, their recommendation could be based solely on the existence of (public) evidence for vaccination (P such that $V(P) = \{NS, MS\}$) and its superior (consensus) certainty compared to the existing evidence against vaccines ($\neg P$).

$$(\varphi \trianglelefteq_e \psi) \rightarrow (\varphi \trianglelefteq_b \psi) \tag{4.1}$$

This example is a trivial simplification of a realistic situation. We are considering very small number of pieces of evidence, a propositional variable instead of a more complex formula, and the model checker we refer to is far from being a practical tool. However, it illustrates the benefit of reasoning about belief and one single dimension that forms that belief—such as certainty of the evidence—within the same logical system. As this example shows, there may be cases where agents can make decisions and provide explanations based on a single, understandable dimension while preserving a connection with a well-established belief measure like Dempster's rule of combination.

This relation between strength of evidence and strength of belief does not hold in general. Because Bel is based on a sum of values, comparing evidence only in terms of its certainty does not explain the outcomes of Bel in all cases—many uncertain evidence will eventually overtake the result of Bel(P) even when P is supported by highly certain evidence. For instance, consider a set of possible states $S = \{a, b, 1, 2, 3, 4, 5, 6, 7, 8\}$ and a quantitative set of uncertain evidence $\boldsymbol{\mathcal{E}}^Q = \{(\{a\}, 0.7), (\{b, 1\}, 0.3), (\{b, 2\}, 0.3), (\{b, 3\}, 0.3), (\{b, 4\}, 0.3), (\{b, 5\}, 0.3), (\{b, 6\}, 0.3), (\{b, 7\}, 0.3), (\{b, 8\}, 0.3)\}$. Then, given P such that $V(P) = \{a\}$ and Q such that $V(Q) = \{b\}$, $Q \trianglelefteq_e P$ but $P \trianglelefteq_b Q$ since $\mathrm{Bel}(P) = 0.12$ and $\mathrm{Bel}(Q) = 0.66$. Therefore, formula (4.1) is not a validity in our logic.

In summary, understanding the relationship between the order on propositions generated by the certainty of evidence and degrees of belief may make computations more transparent and accessible in some cases. Additionally, a similar benefit can be obtained by isolating other features involved in the computation of degrees of belief, such as the consistency among pieces of evidence or the number of pieces of evidence. This example is inspired by (Van der Bles et al., 2019), where the authors present a formal communication of uncertainty based Bayes theorem. Other areas where explicit logical connections between evidence and belief may be useful are administrative processes that must preserve users' information rights or fake news management. In particular, fake news management could benefit from the overtaking effect described earlier in the counterexample against the validity of formula (4.1): if formula (4.1) is invalid, it may indicate an accumulation of uncertain news.

Chapter 5

Computational Complexity of Belief based on Evidence

In theoretical computer science research, expressivity and tractability are often inversely proportional. This means that when a method provides a rich solution from a conceptual standpoint, it often requires considerable computational resources. While sometimes perceived as a limitation, this situation can also encourage further study about the key aspects of the method, in order to discover under which circumstances we can keep its precision without depleting our computational resources. Combining uncertain evidence with the models explored in this thesis is one example of this. Chapter 2 showed that Dempster-Shafer theory, and consequently the multi-layer belief model, adequately represents uncertain evidence and ignorance. However, a straightforward computation of degrees of beliefs by using these methods requires exponential time in general. This chapter is dedicated to delving into the computational complexity of these calculations.

The organization of the chapter is as follows. In Section 5.1, we demonstrate that computing Dempster's rule of combination is a #P-complete problem for basic cases, which admits a fixed-parameter tractable algorithm for a suitable parameter. In Section 5.2, we show that applying topological models of evidence to determine belief in a proposition P given a body of evidence is solvable in polynomial time. Note that in this case evidence is not uncertain. In Section 5.3, we establish that the multi-layer belief model, which extends Dempster's rule of combination, remains #P-complete. Finally, in Section 5.4 we conclude with a concrete example that illustrates the results of the chapter. For details on the theory of computational complexity—and in particular with the complexity class #P—we refer to the textbook by Arora and Barak (2009).

5.1 Applying Dempster's Rule of Combination

In (Orponen, 1990), it is proved that using Dempster's rule of combination is #P-complete. The class #P is a class of counting problems whose equivalent class of decision problems contains the class NP. That is, if P≠NP then Dempster's rule of combination cannot be computed in polynomial time in general (Arora and Barak, 2009). Nevertheless, some results have been established on restricted cases where evidence can be combined in a computationally efficient way (Barnett, 1981; Bergsten and Schubert, 1993; Shafer and Logan, 2008; Shafer et al., 1987). Despite these efforts to understand the opportunities and limitations of applying Dempster's rule of combination, a structured computational complexity analysis has been missing from the literature.

In the first part of this section, we address this gap by performing a structured and detailed complexity analysis of using Dempster's rule of combination. We demonstrate that the problem remains #P-hard when combining basic belief assignments with a single proper focal element each. Additionally, we show that this complexity also remains when the focal elements of the basic belief assignments to be combined are nearly the entire set of possible states. In the second part, we explore the utilization of one of those cases that admit polynomial-time algorithms for combining arbitrary evidence. In particular, we base this exploration on the algorithm proposed by Shafer and Logan (2008) to efficiently compute Dempster's rule of combination over a hierarchical set of evidence. We start our exploration by giving a polynomial-time algorithm for deciding whether an arbitrary body of evidence has a hierarchical structure, enabling the use of Shafer and Logan (2008)'s algorithm. Next, we delve into the problem of extracting hierarchical subsets from the body of evidence. This problem is NP-hard but admits fixed-parameter tractable algorithm—where the parameter is the number of sets to delete. Finally, we generalize the algorithm of Shafer and Logan (2008) to arbitrary sets, demonstrating that this generalized algorithm runs in fixed-parameter tractable time when parameterized by a certain measure of how similar to a hierarchy the body of evidence is.

In the remainder of the section, we will adopt Dempster-Shafer theory terminology to facilitate its reading. Therefore, given a finite set of possible states S, what we called quantitative evidence set $\boldsymbol{\mathcal{E}}^Q = \{(E_1, m_1(E_1)), \dots, (E_\ell, m_\ell(E_\ell))\}$ in the previous chapters, will be now represented as a collection of basic belief assignments $m_1, \dots, m_\ell$ with a single proper focal element $E_1, \dots, E_\ell$, respectively. In addition, if we do not restrict these basic belief assignments to be non-dogmatic, $m_j(E_j)$ may be equal to 1. These cases are not considered in the definition of quantitative evidence set provided in Chapter 2.

Throughout this section, we will also talk about collections of basic belief assignments with two proper focal elements, one being the set complement of the other.

We will call them *dichotomous basic belief assignments*. If we do not specify what kind of input a computational problem accepts, this means that it accepts any collection of basic belief assignments. The inputs of the computational problems considered in this section are collections of basic belief assignments $m_1, \ldots, m_\ell$, with or without further restrictions, that will be referred as *evidence*, *pieces of evidence* or *body of evidence*. These inputs will be merged according to Dempster's rule of combination into $m = \bigoplus_{j=1}^{\ell} m_j$. To differentiate between the basic belief assignments provided as input and the basic belief assignment obtained by Dempster's rule of combination, we will call the latter *combined mass function*. Given a combined mass function m, we assume that the corresponding belief function is computed following Definition 2.5. We refer to it as *belief for* m and denote it by $\mathrm{bel}_m(\cdot)$.

5.1.1 Computational Complexity Under Constraints

Orponen proved that computing beliefs based on the application of Dempster's rule of combination to arbitrary basic belief assignments is #P-hard by giving a reduction from the well-known #SAT problem for conjunctive normal form (CNF) formulas (Orponen, 1990). In this section, we will provide a complexity analysis that builds forth on Orponen's hardness result. In particular, we will identify various restrictions under which the problem remains #P-hard, and we will identify some restrictions that allow polynomial-time algorithms.

The hardness result of Orponen involves basic belief assignments with multiple proper focal elements. So it could be the case that it does not apply to the case where Dempster's rule of combination is only applied to basic belief assignments with a single proper focal element. We begin with showing that, in fact, computing beliefs based on the application of Dempster's rule of combination to this kind of basic belief assignments is #P-hard.

The reduction that we give is very similar to the reduction given by Orponen (1990, Theorem 3.1)—the main difference being that we take a restricted variant of #SAT to reduce from. Nevertheless, we give a description of this reduction to provide the basis for the various complexity results presented in this section that are based on (variations of) this proof. Moreover, our presentation of the correctness argument in the proof differs from that of Orponen, providing the reader with another entry into understanding the argument.

We begin with laying out the precise statements of two computational problems related to the application of Dempster's rule of combination. The difference between these two problems is whether the required output is (1) the combined mass value $m(A)$ or (2) the belief $\mathrm{bel}_m(A)$ for m.

DRC-COMPUTE-MASS
Input: A finite set of possible states S, basic belief assignments $m_1, \dots, m_\ell$ over S, and a set $A \subseteq S$.
Output: $m(A) = (\bigoplus_{j=1}^{\ell} m_j)(A)$.

DRC-COMPUTE-BELIEF
Input: A finite set of possible states S, basic belief assignments $m_1, \dots, m_\ell$ over S, and a set $A \subseteq S$.
Output: $\mathrm{bel}_m(A)$ for $m = \bigoplus_{j=1}^{\ell} m_j$.

Both variants of the problem are #P-hard, even when $m_1, \dots, m_\ell$ only have one proper focal element each.

Theorem 5.1 *DRC-COMPUTE-MASS is #P-complete. Moreover, #P-hardness holds even when restricted to the case where $m_1, \dots, m_\ell$ have a single proper focal element and $|A| = 1$.*

Proof.

We show #P-hardness by providing a reduction from the #P-complete problem #MON-SAT, which concerns counting the number of satisfying truth assignments of a monotone propositional CNF formula (variables occur only positively) (Valiant, 1979). We reiterate that this reduction is entirely similar to the reduction used to show #P-hardness of DRC-COMPUTE-MASS in general (Orponen, 1990)—i.e., without the restriction to one proper focal element per basic belief assignment.

Let $\varphi = c_1 \wedge \dots \wedge c_k$ be a monotone propositional CNF formula over the variables $x_1, \dots, x_\ell$. We define $S = \{1, \dots, k, *\}$ and $A = \{*\}$, and we construct ℓ basic belief assignments $m_1, \dots, m_\ell$ with only one proper focal element each. Each basic belief assignment m_j has as single focal element $T_j = \{*\} \cup \{\ i \mid \text{clause } c_i \text{ does not contain literal } x_j\}$ where:

$$\begin{aligned} &m_j(T_j) = m_j(S) = 1/2, \text{ and} \\ &m_j(B) = 0 \text{ for each } B \in 2^S \setminus \{T_j, S\}. \end{aligned}$$

Now, let $m = \bigoplus_{j=1}^{\ell} m_j$. We will show that $m(A) = m(\{*\}) = q2^{-\ell}$, where q is the number of satisfying truth assignments of φ. Firstly, observe that each m_j assigns non-zero mass only to sets that include $*$, and therefore there is no sequence of sets in the Cartesian product $\boldsymbol{A} = \bigtimes_{j=1}^{\ell} \{T_j, S\}$ that has an empty intersection. As a result, we get that $m(A)$ equals the sum of $\prod_{j=1}^{\ell} m_j(A_j)$ for all sequences $(A_1, \dots, A_\ell) \in \boldsymbol{A}$ such that $\bigcap_{j=1}^{\ell} A_j = \{*\}$. That is, for any sequence $(A_1, \dots, A_\ell) \in \boldsymbol{A}$ such that the intersection of all its elements is $\{*\}$, the product of the masses of its elements will be a summand of $m(A)$.

Moreover, for each sequence $(A_1, \dots, A_\ell) \in \boldsymbol{A}$, it holds that $\prod_{j=1}^{\ell} m_j(A_j) = 2^{-\ell}$.

Consider the following bijection σ between truth assignments $\alpha : \{x_1, \dots, x_\ell\} \to \{0, 1\}$ and sequences $(A_1, \dots, A_\ell) \in \boldsymbol{A}$, where for each α, we let $\sigma(\alpha) = (A_1, \dots, A_\ell)$ such that $A_j = T_j$ if $\alpha(x_j) = 1$, and $A_j = S$ if $\alpha(x_j) = 0$.

It is not difficult to see that for each $\alpha : \{x_1, \dots, x_\ell\} \to \{0, 1\}$ it holds that α satisfies φ if and only if, for $\sigma(\alpha) = (A_1, \dots, A_\ell)$ it holds that $\bigcap_{j=1}^{\ell} A_j = \{*\}$. This suffices to show that $m(A) = q2^{-\ell}$. □

Note that in the proof of Theorem 5.1 we can define the basic belief assignments m_j with only a single proper focal element for the exact reason that φ is monotone. If φ were not monotone, we would have to assign a non-zero mass to the sets $F_j = \{*\} \cup \{ \ i \mid \text{clause } c_i \text{ does not contain literal } \neg x_j\}$ to make the reduction work, as in the original #P-hardness proof for DRC-COMPUTE-MASS (Orponen, 1990). Put differently, due to the fact that φ is monotone, we get that $F_j = \mathrm{S}$.

Corollary 5.1 *DRC-COMPUTE-BELIEF is #P-hard even when restricted to the case where $m_1, \dots, m_\ell$ have a single proper focal element and $|A| = 1$.*

It suffices to notice that for any singleton set A it holds that $m(A) = \mathrm{bel}_m(A)$.

Restricting our attention to basic belief assignments with one proper focal element is not enough to guarantee that we can use Dempster's rule of combination in polynomial time. Now, we will consider two additional restrictions—both based on additional constraints on the size of proper focal elements.

The first additional restriction that we consider is that the size of the proper focal elements of the basic belief assignments are bounded by some fixed constant. This restriction allows us to use Dempster's rule of combination in polynomial time.

Proposition 5.1 *Let $c \in \mathbb{N}$ be a fixed constant. DRC-COMPUTE-MASS and DRC-COMPUTE-BELIEF can be computed in polynomial time if the basic belief assignments $m_1, \dots, m_\ell$ only have proper focal elements of size $\leq c$.*

Proof.

The main idea behind this proof is the following. Whenever you combine two basic belief assignments m_1 and m_2 whose proper focal elements are all of size at most c by using Dempster's rule of combination, the resulting combined mass function assigns positive mass only to sets of size at most c. For any set S of possible states of size n, the number of subsets of size at most c is upper bounded by $(n + 1)^c$—which is a polynomial. Therefore, one can compute the result of Dempster's rule of combination in a brute force fashion in polynomial time. □

Restricting proper focal elements to be of bounded size corresponds to the requirement that all basic belief assignments are highly informative—or in other words, that each basic belief assignment assigns zero mass to all but a few possibilities. Arguably, this occurs only in a very limited set of circumstances, limiting the practical value of the tractability result of Proposition 5.1.

Another restriction, that is perhaps more promising for practical applications, consists in restricting pieces of evidence to be of limited informativeness: allowing only simple basic belief assignments whose proper focal element consists of S with a constant number of possibilities removed. Unfortunately, this restriction does not alleviate the computational intractability of using Dempster's rule of combination.

Proposition 5.2 *Let $c \geq 3$ be a fixed constant. DRC-COMPUTE-MASS and DRC-COMPUTE-BELIEF are #P-hard even when restricted to the case where the basic belief assignments $m_1, \dots, m_\ell$ all have a single proper focal element that is of size $\geq |S| - c$, and where $|A| = 1$.*

Proof.

Similarly to the proof of Theorem 5.1, we adapt the reduction by Orponen (1990). This time, we take as starting point for the reduction the restriction of #MON-SAT where each variable appears in at most 3 clauses. The problem remains #P-complete under this restriction (Greenhill, 2000). The resulting instance then has the property that each of $m_1, \dots, m_\ell$ has a single proper focal element that is of size $\geq |S| - c$, and that $|A| = 1$. □

Nevertheless, as previously mentioned, there are some cases where it is possible to avoid the worst-case computational complexity of using Dempster's rule of combination. We will focus on the algorithm of Shafer and Logan (2008) for efficiently combining hierarchical evidence—i.e., basic belief assignments whose focal elements together form a hierarchy—via Dempster's rule of combination. We will introduce this algorithm, and in Section 5.1.2 we will investigate how to decide if this algorithm can be used to efficiently aggregate (a subset of) the evidence in a given situation. Moreover, we will study a way to extend this algorithm to arbitrary sets of evidence.

Definition 5.1 (Hierarchy). Let S be a finite set of possible states. A set $\mathcal{H} = \{A_1, \dots, A_\ell\}$ of focal elements $A_j \subseteq S$ is a *hierarchy* over S if there exists a tree where the root node is labelled with S, and all other nodes are labelled with an element $A_j \in \mathcal{H}$ such that: (1) if a node labelled with A_k is the child of a node labelled with A_j, then $A_k \subseteq A_j$, and (2) if two nodes labelled with A_j and A_k are siblings, then $A_j \cap A_k = \emptyset$. In other words, a set $\mathcal{H} = \{A_1, \ldots, A_\ell\} \subseteq 2^S$ is a hierarchy if and only if for all A_j and $A_k \in \mathcal{H}$ it holds that: if $A_j \cap A_k \neq \emptyset$, then $A_k \subseteq A_j$ or $A_j \subseteq A_k$.

Example 5.1 Hierarchy

Consider $S = \{a, b, c, d, e\}$. Then $\mathcal{H} = \{A_1, \ldots, A_6\}$, for $A_1 = \{a, b, c\}$, $A_2 = \{d, e\}$, $A_3 = \{a, b\}$, $A_4 = \{a\}$, $A_5 = \{e\}$ and $A_6 = \{d\}$, is a hierarchy.

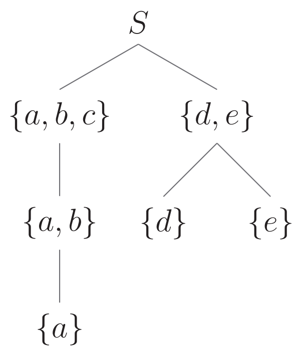


Figure 5.1: Diagram representation of hierarchy $\mathcal{H} = \{A_1, \ldots, A_6\}$.

Theorem 5.2 (Shafer and Logan, 2008). *Given a hierarchy $\mathcal{H}$ over a set of possible states S and a collection of basic belief assignments $m_1, \ldots, m_\ell$ with one single proper focal element each such that it is either an element of $\mathcal{H}$ or the set complement of one, then $\text{bel}_m(A)$, $\text{bel}_m(\overline{A})$, $\text{plau}_m(A)$ and $\text{plau}_m(\overline{A})$, for $m = \bigoplus_{j=1}^{\ell} m_j$, can be computed in polynomial time for all $A \in \mathcal{H}$.*

The algorithm provided by Theorem 5.2 does not allow us to efficiently compute the belief (or plausibility) of arbitrary sets $A \subseteq S$, but only of sets $A \in \mathcal{H}$ (or their complements). Nevertheless, it is plausible to assume that if the available evidence in a given application domain forms a hierarchy (as specified in Definition 5.1), then sets $P \subseteq S$ such that we are interested in knowing their degree of belief also match this hierarchy. This is the case, for example, in the setting of diagnostic reasoning in medicine, where hierarchical evidence naturally appears (Gordon and Shortliffe, 2008).

5.1.2 Computational Complexity Identifying Hierarchies

In this subsection, we explore three possible approaches to take advantage of the computational gain of applying Shafer and Logan (2008)'s algorithm for hierarchical evidence. First, we prove that deciding whether an arbitrary body of evidence forms a hierarchy is a tractable problem. We call this approach *plac-*

ing all evidence in a hierarchy. Second, we study the problem of extracting a hierarchy from an arbitrary body of evidence. We show that this problem is fixed-parameter tractable for a suitable parameter. This part of the discussion is titled *placing as much evidence as possible in a hierarchy.* Lastly, we propose a fixed-parameter tractable algorithm to combine evidence according to Dempster's rule of combination. Our algorithm uses Shafer and Logan (2008)'s algorithm in some intermediate steps. This result appears under the name of *using hierarchical structure in arbitrary bodies of evidence.*

Placing All Evidence in a Hierarchy

Suppose that a researcher has performed different experiments that provide evidence for various subsets of a set of possible states S and is interested in knowing which elements of S have the most support according to it. It would be very useful to know whether these focal elements form a hierarchy, in which case they could use the Shafer-Logan algorithm (Theorem 5.2) to efficiently compute beliefs based on the combination of all the evidence. This algorithm works for the case where for each set A in the hierarchy a dichotomous basic belief assignment is given—which assigns weight to A and to its complement $S \setminus A$. Therefore, we take as starting point for this problem a set $\mathcal{A}$ with ℓ pairs of sets, each consisting of a focal element and its set complement—one per each dichotomous basic belief assignment. We will give a polynomial-time algorithm that decides, given a set $\mathcal{A}$ of such pairs of sets, whether we can form a hierarchy $\mathcal{H}$ by taking exactly one set from each pair—in which case we say that $\mathcal{A}$ *admits a hierarchy*—and that computes what this hierarchy would look like.

Definition 5.2 (Conflict). We will say that there exists a *conflict between two focal elements A_j and A_k* if $A_j \not\subseteq A_k$, $A_k \not\subseteq A_j$, and $A_j \cap A_k \neq \emptyset$.

We will denote such a conflict with $A_j \rightleftharpoons A_k$.

Theorem 5.3 *Let S be a finite set of possible states. Given a set $\mathcal{A} = \{(B_j, \overline{B}_j)\}_{j=1}^{\ell}$ of pairs, where for each $1 \leq j \leq \ell$, B_j and $\overline{B}_j = S \setminus B_j$ are complementary sets over S, the following are equivalent:*

(a) There exists a hierarchy $\mathcal{H}$ consisting of exactly one set from each pair in $\mathcal{A}$.

(b) There exists a set $\{A_1, \ldots, A_\ell\}$ of focal elements formed by exactly one element of each pair in $\mathcal{A}$ such that for each two distinct $A_j, A_k \in \{A_1, \ldots, A_\ell\}$ it holds that $A_j \not\rightleftharpoons A_k$.

(c) The following conjunctive normal form formula with disjunctions of two literals (2CNF formula) $\varphi = \varphi_1 \wedge \varphi_2$ over the variables $x_1, \ldots, x_\ell, y_1, \ldots, y_\ell$ is satisfiable. For each $1 \leq j \leq \ell$, let $\nu(B_j) = x_j$ and $\nu(\overline{B}_j) = y_j$. Then φ_1

consists of the clauses $(\nu(B_j) \vee \nu(\overline{B}_j))$ *and* $(\neg\nu(B_j) \vee \neg\nu(\overline{B}_j))$ *for each* $1 \leq j \leq \ell$*, and* φ_2 *consists of the clauses* $(\neg\nu(A_j) \vee \neg\nu(A_k))$ *for each* $A_j \in \{B_j, \overline{B}_j\}$ *and* $A_k \in \{B_k, \overline{B}_k\}$ *such that* $A_j \rightleftharpoons A_k$*.*

Proof.

One can straightforwardly show that (a) and (b) are equivalent by using Definitions 5.1 and 5.2.

We then show that (b) implies (c). Suppose that there is a set $\{A_1, \ldots, A_\ell\}$ of focal elements formed by exactly one element of each pair in $\mathcal{A}$ such that for each two distinct $A_j, A_k \in \{A_1, \ldots, A_\ell\}$ it holds that $A_j \not\rightleftharpoons A_k$. We then define a truth assignment α that satisfies φ. For each $1 \leq j \leq \ell$, let $\alpha(x_j) = 1$ and $\alpha(y_j) = 0$ if $B_j \in \{A_1, \ldots, A_\ell\}$ and let $\alpha(x_j) = 0$ and $\alpha(y_j) = 1$ if $B_j \notin \{A_1, \ldots, A_\ell\}$. This assignment satisfies φ_1 because for each $1 \leq j \leq \ell$ there is exactly one of $B_j, \overline{B}_j$ in the set $\{A_1, \ldots, A_\ell\}$. The clauses in φ_2 are also satisfied by α because there are no two distinct $A_j, A_k \in \{A_1, \ldots, A_\ell\}$ such that $A_j \rightleftharpoons A_k$.

Finally, we show that (c) implies (b). Take a truth assignment α that satisfies φ. For each $1 \leq j \leq \ell$, we let $A_j = B_j$ if $\alpha(x_j) = 1$ and $A_j = \overline{B}_j$ if $\alpha(x_j) = 0$. Then $\{A_1, \ldots, A_\ell\}$ contains exactly one element of each pair in $\mathcal{A}$. Now, to derive a contradiction, suppose that there were two distinct $A_j, A_k \in \{A_1, \ldots, A_\ell\}$ with $A_j \rightleftharpoons A_k$. Then φ would contain the clause $(\neg\nu(A_j) \vee \neg\nu(A_k))$, and $\alpha(\nu(A_j)) = \alpha(\nu(A_k)) = 1$, and so α would not satisfy φ, which contradicts our assumption. Therefore, we can conclude that there are no two distinct $A_j, A_k \in \{A_1, \ldots, A_\ell\}$ such that $A_j \rightleftharpoons A_k$. □

In fact, the proof of Theorem 5.3 also shows that we can efficiently construct a hierarchy $\mathcal{H}$ containing exactly one element from each pair in $\mathcal{A}$ from a truth assignment for the 2CNF formula.

Corollary 5.2 *Let* S *be a finite set of possible states and* $\mathcal{A}$ *be a set of focal elements closed under complement. Moreover, let* φ *be the 2CNF formula described in Theorem 5.3. Then from any assignment* α *that satisfies* φ*, we can in polynomial time construct a hierarchy* $\mathcal{H}$ *containing exactly one element from each pair in* $\mathcal{A}$*.*

Since one can in linear time decide whether a given 2CNF formula is satisfiable (and if so, find a satisfying truth assignment) (Aspvall et al., 1979), we can in polynomial time decide whether the pairs in $\mathcal{A}$ admit a hierarchy, and compute such a hierarchy if this is the case.

Placing As Much Evidence As Possible in a Hierarchy

Even if deciding whether a given set of evidence can all be placed in a single hierarchy can be efficiently done, placing the evidence in an hierarchy is not always possible. Therefore, we will study algorithms to form hierarchies that accommodate much (but not all) of a given set $\mathcal{A}$ of evidence. Since we suppose that we are given a set $\mathcal{A}$ with ℓ pairs $(B_j, \overline{B}_j)$ of sets—each consisting of a focal element and its complement—if there is no single hierarchy in line with all of this evidence, we can only obtain a hierarchy by selecting one set from some (but not all) pairs $(B_j, \overline{B}_j)$.

One way to make such a selection is the following. We take the 2CNF formula φ from Theorem 5.3, and adapt it into $\varphi' = \varphi_1' \wedge \varphi_2$, where φ_1' consists only of the clauses $(\neg\nu(B_j) \vee \neg\nu(\overline{B}_j))$. Then, any satisfying truth assignment for φ' corresponds to a hierarchy that fits a subset of the evidence. However, by taking this crude approach, we have no influence on how much of the evidence is accommodated by the resulting hierarchy—for example, one can satisfy φ' by setting all variables to false, which corresponds to the trivial, empty hierarchy.

We will start with distinguishing some structure in the set of conflicts between focal elements that will turn out to be useful to develop algorithms for finding (large) partial hierarchies. In the remainder, we will assume that a set $\mathcal{A}$ with ℓ pairs $P_j = (B_j, \overline{B}_j)$ of complementary sets is given.

Definition 5.3 (Conflict between pairs of focal elements). Let $P_1 = (A_1, \overline{A}_1)$ and $P_2 = (A_2, \overline{A}_2)$ be two different pairs of complementary sets. Moreover, let r be the number of conflicts between the sets appearing in P_1 and P_2—that is, $r = |\{(B_1, B_2) \mid B_1 \in \{A_1, \overline{A}_1\}, B_2 \in \{A_2, \overline{A}_2\}, B_1 \rightleftharpoons B_2\}|$. We then say that there are r conflicts between P_1 and P_2, and we denote this by $P_1 \overset{r}{\rightleftharpoons} P_2$. We write $P_1 \rightleftharpoons P_2$ if $P_1 \overset{r}{\rightleftharpoons} P_2$ for some $r > 0$. If $P_1 \overset{1}{\rightleftharpoons} P_2$, we say that there is a *single conflict* and if $P_1 \overset{4}{\rightleftharpoons} P_2$, we say that there is a *total conflict* between P_1 and P_2.

Moreover, we define $\mathcal{C}^{\mathcal{A}} = \{(P_j, P_k) \mid P_j, P_k \in \mathcal{A}, P_j \rightleftharpoons P_k\}$ and $\mathcal{C}_r^{\mathcal{A}} = \{(P_j, P_k) \mid P_j, P_k \in \mathcal{A}, P_j \overset{r}{\rightleftharpoons} P_k\}$.

By establishing that between any two pairs P_1 and P_2 of complementary focal elements, there is either a single conflict or there is a total conflict, we can characterize the existence of a hierarchy in terms of single conflicts.

Lemma 5.1 *For each set of possible states S and each set $\mathcal{A}$ of pairs of complementary focal elements, $\mathcal{C}^{\mathcal{A}} = \mathcal{C}_1^{\mathcal{A}} \cup \mathcal{C}_4^{\mathcal{A}}$.*

Proof.

Let us see that two distinct pairs $P_j = (B_j, \overline{B}_j)$ and $P_k = (B_k, \overline{B}_k)$ of $\mathcal{A}$ have at least one conflict. If $P_j \not\rightleftharpoons P_k$ then $B_j \not\rightleftharpoons B_k$ (1), $\overline{B}_j \not\rightleftharpoons B_k$ (2), $B_j \not\rightleftharpoons \overline{B}_k$ (3) and $\overline{B}_j \not\rightleftharpoons \overline{B}_k$ (4). For (1), at least one of these three conditions must hold:

(a) $B_j \subset B_k$,

(b) $B_k \subset B_j$ or

(c) $B_j \cap B_k = \emptyset$.

If $B_j \subset B_k$, then $B_k \cap \overline{B}_j \neq \emptyset$ since the inclusion is strict. In addition, $B_k \not\subset \overline{B}_j$ and, if $B_k \neq S$, $\overline{B}_j \not\subset B_k$. Therefore, $\overline{B}_j \rightleftharpoons B_k$, which contradicts (2).

A similar reasoning can show that if $B_k \subset B_j$, and $B_j \neq S$, then $B_j \rightleftharpoons \overline{B}_k$, contradicting (3).

Finally, if $B_j \cap B_k = \emptyset$, then $B_k \subset \overline{B}_j$ and $B_j \subset \overline{B}_k$, so $\overline{B}_j \not\subset \overline{B}_k$ and $\overline{B}_k \not\subset \overline{B}_j$ respectively. Furthermore, as these inclusions are strict, $\overline{B}_j \cap \overline{B}_k \neq \emptyset$. This means that $\overline{B}_j \rightleftharpoons \overline{B}_k$ and contradicts (4).

Due to all of the above three conditions implies a contradiction, we can conclude that there is at least one conflict between elements of P_j and P_k.

Now, let us prove that if there is a conflict between B_j, B_k and $((B_j, \overline{B}_j), (B_k, \overline{B}_k)) \notin \mathcal{C}_4^{\mathcal{A}}$ then $\overline{B}_j \cap \overline{B}_k = \emptyset$, $\overline{B}_j \subset B_k$ and $\overline{B}_k \subset B_j$, and as a consequence, $((B_j, \overline{B}_j), (B_k, \overline{B}_k)) \in C_1$.

On the one hand, $\overline{B}_j \cap \overline{B}_k = \emptyset$ implies $\overline{B}_j \subset B_k$ and $\overline{B}_k \subset B_j$, since that empty intersection implies that all the elements of $\overline{B}_j$ (respectively. $\overline{B}_k$) are contained in the complement of $\overline{B}_k$ (respectively $\overline{B}_j$).

On the other hand, if $\overline{B}_j \cap \overline{B}_k \neq \emptyset$, then not only B_j has a conflict with B_k but also (a) B_j has a conflict with $\overline{B}_k$, (b) $\overline{B}_j$ has a conflict with B_k and (c) $\overline{B}_j$ has a conflict with $\overline{B}_k$.

(a) First, $B_j \rightleftharpoons B_k$ implies $B_j \not\subset B_k$, so there is an element in B_j which belong to $\overline{B}_k$ and $B_j \cap \overline{B}_k \neq \emptyset$. Secondly, $B_j \cap B_k \neq \emptyset$ so $B_j \not\subset \overline{B}_k$. Finally, $\overline{B}_j \cap \overline{B}_k \neq \emptyset$, so there is an element in $\overline{B}_k$ which is not in B_j, i.e., $\overline{B}_k \not\subset B_j$.

(b) The conflict $B_j \rightleftharpoons B_k$ also implies $B_j \cap B_k \neq \emptyset$ so $B_k \not\subset \overline{B}_j$. In addition, $\overline{B}_j \cap \overline{B}_k \neq \emptyset$ proves that $\overline{B}_j \not\subset B_k$. Lastly, if $\overline{B}_j \cap B_k = \emptyset$ then $B_k \subset B_j$ which is not possible since $B_j \rightleftharpoons B_k$.

(c) Since our hypotheses is that $\overline{B}_j \cap \overline{B}_k \neq \emptyset$, $\overline{B}_j \not\subset \overline{B}_k$ and $\overline{B}_k \not\subset \overline{B}_j$ for $B_k \not\subset B_j$ and $B_j \not\subset B_k$ respectively.

Therefore, if $B_j \rightleftharpoons B_k$ then $(B_j, \overline{B}_j) \rightleftharpoons_{1} (B_k, \overline{B}_k)$ or $(B_j, \overline{B}_j) \rightleftharpoons_{4} (B_k, \overline{B}_k)$. $\square$

Proposition 5.3 *Let S be a finite set of possible states, and let $\mathcal{A}$ be a set with pairs $P_j = (B_j, \overline{B}_j)$ of complementary sets over S. Then there exists a hierarchy containing exactly one of B_j and $\overline{B}_j$ for each pair P_j if and only if $\mathcal{C}^{\mathcal{A}} = \mathcal{C}_1^{\mathcal{A}}$.*

Proof.

Firstly, suppose that $\mathcal{C}^{\mathcal{A}} = \mathcal{C}_1^{\mathcal{A}}$. We claim that the formula φ from Theorem 5.3 is satisfied by the following truth assignment α, which suffices to show the existence of a suitable hierarchy. For each j, let $\alpha(x_j) = 1$ and $\alpha(y_j) = 0$ if $|B_j| \leq |\overline{B}_j|$, and let $\alpha(x_j) = 0$ and $\alpha(y_j) = 1$ otherwise.

Conversely, suppose that there exists a suitable hierarchy. Then the 2CNF formula φ from Theorem 5.3 is satisfiable. We show that no pair (P_j, P_k) can belong to $\mathcal{C}_4^{\mathcal{A}}$. To derive a contradiction, suppose that there are $P_j = (B_j, \overline{B}_j)$ and $P_k = (B_k, \overline{B}_k)$ such that $(P_j, P_k) \in \mathcal{C}_4^{\mathcal{A}}$. Then, by construction, φ would contain the following clauses: $(\nu(B_j) \vee \nu(\overline{B}_j))$, $(\neg\nu(B_j) \vee \neg\nu(\overline{B}_j))$, $(\nu(B_k) \vee \nu(\overline{B}_k))$, $(\neg\nu(B_k) \vee \neg\nu(\overline{B}_k))$, $(\neg\nu(B_j) \vee \neg\nu(B_k))$, $(\neg\nu(B_j) \vee \neg\nu(\overline{B}_k))$, $(\neg\nu(B_k) \vee \neg\nu(B_j))$, and $(\neg\nu(B_k) \vee \neg\nu(\overline{B}_j))$. Thus φ would be unsatisfiable, which contradicts Theorem 5.3. By Lemma 5.3, we then know that $\mathcal{C}^{\mathcal{A}} = \mathcal{C}_1^{\mathcal{A}}$. $\square$

Now, we study the problem of finding partial hierarchies—among a given set $\mathcal{A}$ of pairs of complementary focal elements—that are as large as possible. In particular, we will show that this problem is closely related to the classical problem VERTEX-COVER, that consists in deciding if a graph has a vertex cover of a given size. Concretely, we show that there is a polynomial-time reduction from VERTEX-COVER to the problem of finding a partial hierarchy of a given size—showing that the latter problem is NP-hard. We also show that there is a polynomial-time reduction in the other direction—allowing fixed-parameter tractable algorithms for VERTEX-COVER to be employed for finding hierarchies.

We start with giving a formal definition for the decision problem of finding large partial hierarchies, and showing that this problem is NP-complete.

PARTIAL-HIERARCHY

Input: A set S of possible states, a set $\mathcal{A} = \{(B_j, \overline{B}_j)\}_{j=1}^{\ell}$ of complementary pairs of focal elements over S, and a positive integer $r \in \mathbb{N}$.

Question: Is there a hierarchy $\mathcal{H} \subseteq \{B_j, \overline{B}_j \mid 1 \leq j \leq \ell\}$ of size at least r, such that $\mathcal{H} \cap \{B_j, \overline{B}_j\} \leq 1$ for each j?

Theorem 5.4 *PARTIAL-HIERARCHY is NP-complete.*

Proof.

It is straightforward to show that the problem is in NP, and we omit further details on this. To show NP-hardness, we give a reduction from VERTEX-COVER. Let $G = (V, E)$ be an undirected graph with $V = \{v_1, \ldots, v_\ell\}$, and let $t \in \mathbb{N}$. We construct an instance of PARTIAL-HIERARCHY. We let $S = \{*\} \cup V \cup E$. Moreover, we define $\mathcal{A} = \{(B_j, \overline{B}_j)\}_{j=1}^{\ell}$ by letting $B_j = \{v_j\} \cup \{e \in E \mid v_j \in e\}$ and $\overline{B}_j = S \setminus B_j$. Finally, we let $r = \ell - t$.

Then G has a vertex cover—i.e., a set of vertices such that every edge in the graph is incident to at least one vertex in the set—of size t if and only if there is a hierarchy $\mathcal{H} \subseteq \{B_j, \overline{B}_j \mid 1 \leq j \leq \ell\}$ of size r. In particular, for any $C \subseteq V$ it holds that C is a vertex cover of G if and only if $\mathcal{A}_C = \{(B_j, \overline{B}_j) \mid v_j \in V \setminus C\}$ admits a hierarchy (in the sense of Theorem 5.3).

To show this, the following claim is central. For each $v_j, v_k \in V$ such that $j \neq k$, if $\{v_j, v_k\} \in E$, then $(B_j, \overline{B}_j) \mathrel{\lesssim 4 \gtrsim} (B_k, \overline{B}_k)$, and if $\{v_j, v_k\} \notin E$, then $(B_j, \overline{B}_j) \mathrel{\lesssim 1 \gtrsim} (B_k, \overline{B}_j)$. The above correspondence between vertex covers C of G and partial hierarchies $\mathcal{A}_C$ is straightforward to show, using this claim and Proposition 5.3. □

Proposition 5.4 *There is a polynomial-time reduction from PARTIAL-HIERARCHY to VERTEX-COVER that maps instances $(S, \mathcal{A}, t)$ to instances with $r = |\mathcal{A}| - t$.*

Proof.

We describe the main lines of this reduction, and we omit a proof of correctness—which is analogous to the proof of Theorem 5.4. Let S be a finite set of possible states, $\mathcal{A} = \{(B_j, \overline{B}_j)\}_{j=1}^{l}$ a set of complementary pairs over S, and t a positive integer. We construct $G = (V, E)$ by letting $V = \{v_1, \ldots, v_\ell\}$ and $E = \{\{v_j, v_k\} \mid (B_j, \overline{B}_j) \mathrel{\lesssim 4 \gtrsim} (B_k, \overline{B}_k)\}$. Then $\mathcal{A}$ and r form a yes-instance for PARTIAL-HIERARCHY if and only if G has a vertex cover of size $r = \ell - t$, and solutions are in one-to-one correspondence. □

The result of Proposition 5.4 shows that we can use fixed-parameter tractable algorithms for VERTEX-COVER to find partial hierarchies, and that such algorithms can be expected to run efficiently in cases where we can obtain a hierarchy from $\mathcal{A}$ by removing only few items. In particular, considering a set of possible states of size n, we can find vertex covers of size r in time $O(1.2738^r + rn)$ (Chen et al., 2010), which is a running time that is manageable whenever $r = \ell - t$ is reasonably small. Additionally, one could employ approximation algorithms for finding minimum-size vertex covers–see, e.g., (Arora and Barak, 2009)–to get partial hierarchies that approximate those of maximum size. This line of work points to several interesting open problems. One example is to study how filtering the original set of evidence–e.g., to get a hierarchy structure–affects the degree of belief in a proposition when applying Dempster's rule of combination.

Using Hierarchical Structure in Arbitrary Bodies of Evidence

We have studied the problem of determining in what situations–possibly after disregarding some pieces of evidence–evidence is entirely aligned with a hierarchy. Of course, there are also situations where it is not possible to obtain a hierarchy from a body of evidence–and where disregarding evidence is undesirable or inappropriate. Now, we take some initial steps towards algorithmically using hierarchical structure to combine evidence also in cases where the evidence is not entirely in line with any hierarchy.

Specifically, we first introduce a measure of how much (of a particular type of) hierarchical structure there is in any set $\mathcal{A}$ of focal elements. We then provide an algorithm to compute the belief function corresponding to the combination of a given evidence set–based on applying Dempster's rule of combination to basic belief assignments with a single proper focal element–that works efficiently when there is a high degree of hierarchical structure in $\mathcal{A}$.

The main idea behind this measure and the algorithm is as follows. Whenever there are focal elements $A_1, A_2 \in \mathcal{A}$ that are conflicting—in the sense of Definition 5.3—we merge them together. We do this merging iteratively until there are no conflicts remaining, and thus until we have a hierarchy. The algorithm, roughly, works in a two-step fashion: (i) for all focal elements in the hierarchy that are the result of such a merging operation, we compute the combined belief using a brute-force algorithm; (ii) we use the algorithm of Theorem 5.2 to combine these intermediate results—that involve the previously merged focal elements—with the evidence for focal elements that are not the result of any merging operation.

Step (i) of this algorithm takes exponential time, but this is only exponential in the size of the merged focal elements. As the measure of the amount of hierarchical structure in the set $\mathcal{A}$ of focal elements we take the size r of the largest focal element resulting from this iterative merging process. The smaller this number r, the more hierarchical structure the set $\mathcal{A}$ contains, and in fact, if $\mathcal{A}$ is already a hierarchy, then $r = 0$. The running time of the algorithm then is $2^r \cdot \text{poly}(|x|)$, where x denotes the size of the problem input. In other words, the algorithm runs in *fixed-parameter tractable time*, when we consider as parameter the amount of hierarchical structure.

Let us now work out this idea in more detail, and let us begin by introducing the measure of the amount of hierarchical structure in any given set of focal elements.

Definition 5.4 (Corresponding merged hierarchy). Let S be a finite set of possible states and $\mathcal{A} = \{A_1, \ldots, A_\ell\}$ a set of focal elements $A_j \subseteq$ S. Then the *merged hierarchy $\mathcal{H}_\mathcal{A}$ corresponding to $\mathcal{A}$* is defined by the following procedure. Initially, let $\mathcal{A}_{\text{origin}} = \mathcal{A}$ and $\mathcal{A}_{\text{merged}} = \emptyset$, and then iteratively update $(\mathcal{A}_{\text{origin}}, \mathcal{A}_{\text{merged}})$ using the following rules until the rules no longer apply.

- If there are $A_j, A_k \in \mathcal{A}_{\text{origin}}$ such that both (i) $A_j \cap A_k \neq \emptyset$ and (ii) neither $A_j \subseteq A_k$ nor $A_k \subseteq A_j$, then replace $\mathcal{A}_{\text{origin}}$ by $\mathcal{A}_{\text{origin}} \setminus \{A_j, A_k\}$ and add $A_j \cup A_k$ to $\mathcal{A}_{\text{merged}}$.
- If there is some $A_j \in \mathcal{A}_{\text{origin}}$ and some $A_k \in \mathcal{A}_{\text{merged}}$ such that both (i) $A_j \cap A_k \neq \emptyset$ and (ii) $A_k \not\subseteq A_j$, then replace $\mathcal{A}_{\text{origin}}$ by $\mathcal{A}_{\text{origin}} \setminus \{A_j\}$ and replace $\mathcal{A}_{\text{merged}}$ by $\mathcal{A}_{\text{merged}} \setminus \{A_k\} \cup \{A_j \cup A_k\}$.
- If there are $A_j, A_k \in \mathcal{A}_{\text{merged}}$ such that $A_j \cap A_k \neq \emptyset$, then replace $\mathcal{A}_{\text{merged}}$ by $(\mathcal{A}_{\text{merged}} \setminus \{A_j, A_k\}) \cup \{A_j \cup A_k\}$.

Finally, let $\mathcal{H}_\mathcal{A} = \mathcal{A}_{\text{origin}} \cup \mathcal{A}_{\text{merged}}$.

If $\mathcal{A}$ is already a hierarchy, then none of these rules ever applies, and thus $\mathcal{H}_\mathcal{A} = \mathcal{A}$. No matter in which order you apply the rules in this iterative procedure, the result does not change. In other words, for any $\mathcal{A}$, the hierarchy $\mathcal{H}_\mathcal{A}$ is uniquely defined.

Proposition 5.5 *For each set $\mathcal{A}$ of focal elements, the procedure in Definition 5.4 yields a unique $\mathcal{H}_\mathcal{A}$, regardless of the order in which rules are applied. Moreover, $\mathcal{H}_\mathcal{A}$ is a hierarchy, and for each $A \in \mathcal{A}$ there is some $H \in \mathcal{H}_\mathcal{A}$ such that $A \subseteq H$.*

Proof.

Each of the rules only merges sets, which directly gives us termination and the property that for each $A \in \mathcal{A}$ there is some $H \in \mathcal{H}_\mathcal{A}$ such that $A \subseteq H$. If the resulting $\mathcal{H}_\mathcal{A}$ were not a hierarchy, then one could still apply a rule, which proves that $\mathcal{H}_\mathcal{A}$ must be a hierarchy. Uniqueness can be proved with the observation that the effects of the rules only strictly increase the sets in $\mathcal{A}_{\text{merged}}$, and the preconditions of the rules are monotone—in the sense that making sets in $\mathcal{A}_{\text{merged}}$ larger will not make a previously applicable rule not applicable anymore. □

Having the notion of corresponding merged hierarchies in place, we introduce the level of merging needed to construct $\mathcal{H}_\mathcal{A}$ from $\mathcal{A}$ as a way to measure the amount of hierarchical structure in $\mathcal{A}$.

Definition 5.5 (Level of merging). Let S be a finite set of possible states and $\mathcal{A} = \{A_1, \ldots, A_\ell\}$ a set of focal elements $A_j \subseteq S$. We define the *level of merging* needed to construct the merged hierarchy $\mathcal{H}_\mathcal{A}$ from $\mathcal{A}$ to be $r = \max_{A \in \mathcal{A}_{\text{merged}}} |A|$, where $\mathcal{A}_{\text{origin}}$ and $\mathcal{A}_{\text{merged}}$ are given by the procedure described in Definition 5.4.

The procedure described in Definition 5.4 gives us a polynomial-time algorithm to compute both $\mathcal{H}_{\mathcal{A}}$ and r.

Proposition 5.6 *For each $\mathcal{A}$, we can in polynomial time compute its corresponding hierarchy $\mathcal{H}_{\mathcal{A}}$ and compute the level r of merging needed to construct $\mathcal{H}_{\mathcal{A}}$ from $\mathcal{A}$.*

Proof.

The procedure described in Definition 5.4 terminates in polynomial time and produces $\mathcal{H}_{\mathcal{A}}$ and r. □

Example 5.2 Merged Hierarchy

Consider $S = \{a, b, c, d, e\}$ and $\mathcal{A} = \{\{a\}, \{a, b\}, \{b, c\}, \{a, b, c, d\}, \{d\}, \{e\}\}$. Then $\mathcal{H}_{\mathcal{A}} = \{\{a, b, c\}, \{a, b, c, d\}, \{d\}, \{e\}\}$ and the level r of merging needed to construct $\mathcal{H}_{\mathcal{A}}$ from $\mathcal{A}$ is 3, as $\{a, b, c\}$ is the largest element in the set $\mathcal{A}_{\text{merged}}$ resulting from the procedure described in Definition 5.4.

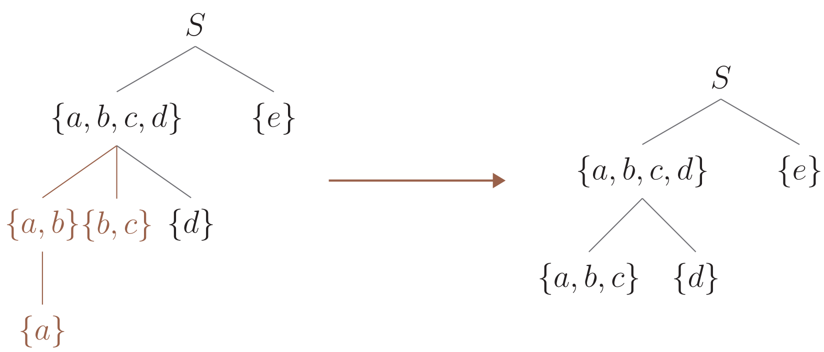


Figure 5.2: Diagram representation of set $\mathcal{A} = \{\{a\}, \{a, b\}, \{b, c\}, \{d\}, \{e\}, \{a, b, c, d\}\}$ (left) and the hierarchy $\mathcal{H}_{\mathcal{A}}$. The colored arrow represents the process of constructing $\mathcal{H}_{\mathcal{A}}$ from $\mathcal{A}$. The colored sets in $\mathcal{A}$ are the ones merged during the process.

We now have everything in place to present our fixed-parameter tractable algorithm that extends the result of Shafer and Logan (2008) (Theorem 5.2) to arbitrary sets $\mathcal{A}$ of focal elements. We precede it with two lemmas necessary for its proof. They can be established straightforwardly using the definition of Dempster's rule of combination.

Lemma 5.2 *Let S be a finite set of possible states, let m_1, m_2, m_3 be basic belief assignments over S, and let $A \subseteq S$ be such that* $\mathrm{bel}_{m_1}(A) + m_1(S) = 1$, $\mathrm{bel}_{m_2}(A) + m_2(S) = 1$, *and* $\mathrm{bel}_{m_3}(S \setminus A) + m_3(S) = 1$. *Moreover, let $m_{1,3} = m_1 \oplus m_3$ and let $m_{2,3} = m_2 \oplus m_3$. Then* $\mathrm{bel}_{m_{1,3}}(A) = \mathrm{bel}_{m_{2,3}}(A)$ *and* $\mathrm{bel}_{m_{1,3}}(B) = \mathrm{bel}_{m_{2,3}}(B)$ *for each $B \subseteq S \setminus A$.*

Lemma 5.3 *Let S be a finite set of possible states, let m_1, m_2 be basic belief assignments over S, and let $A \subseteq S$ be such that* $\mathrm{bel}_{m_1}(A) + m_1(S) = 1$, *and* $\mathrm{bel}_{m_2}(S \setminus A) + m_2(S) = 1$. *Then* $\mathrm{bel}_{m_1}(B)/\mathrm{bel}_{m_1}(A) = \mathrm{bel}_{m_{1,2}}(B)/\mathrm{bel}_{m_{1,2}}(A)$ *for each $B \subseteq A$, where $m_{1,2} = m_1 \oplus m_2$.*

Theorem 5.5 *Let S be a finite set of possible states and let $\mathcal{A} = \{A_1, \dots, A_\ell\}$ be a set of focal elements $A_j \subseteq S$. Moreover, let $\mathcal{H}_\mathcal{A}$ be the hierarchy corresponding to $\mathcal{A}$, and let r be the level of merging needed to construct $\mathcal{H}_\mathcal{A}$ from $\mathcal{A}$. Then, given some basic belief assignments $m_1, \dots, m_\ell$, with one single proper focal element each, and assuming these focal elements are elements of $\mathcal{A}$, for each $A \in \mathcal{A}$ we can compute* $\mathrm{bel}_m(A)$, $\mathrm{bel}_m(\overline{A})$, $\mathrm{plau}_m(A)$ *and* $\mathrm{plau}_m(\overline{A})$, *for $m = \bigoplus_{j=1}^{\ell} m_j$, in time $2^r \cdot$ poly$(|x|)$ where x denotes the problem input.*

Proof.

We describe how to compute $\mathrm{bel}_m(A)$. This procedure can be straightforwardly modified to compute $\mathrm{bel}_m(\overline{A})$, $\mathrm{plau}_m(A)$ and $\mathrm{plau}_m(\overline{A})$ as well. We may assume without loss of generality that, for each $A \in \mathcal{A}$, there is exactly one basic belief assignment among $m_1, \dots, m_\ell$ that has A as proper focal element–call this basic belief assignment m_A.

We will use the following procedure. Firstly, we construct $\mathcal{H}_\mathcal{A}$, together with $\mathcal{A}_{\text{origin}}$ and $\mathcal{A}_{\text{merged}}$, as described in Definition 5.4. Then, for each $A \in \mathcal{A}_{\text{merged}}$, we use a brute-force approach to compute $m_A = \bigoplus_{A' \in \mathcal{A}, A' \subseteq A} m_{A'}$, and we construct the basic belief assignments m'_A with one proper focal element A such that $m'_A(A) = \mathrm{bel}_{m_A}(A)$. Then we use Theorem 5.2, using m_A for each $A \in \mathcal{A}_{\text{origin}}$ and m'_A for each $A \in \mathcal{A}_{\text{merged}}$, to compute $\mathrm{bel}_m(H)$ for each $H \in \mathcal{H}_\mathcal{A}$. By Lemma 5.2, we can safely replace m_A by m'_A in this computation, for each $A \in \mathcal{A}_{\text{merged}}$. What remains is to compute $\mathrm{bel}_m(A)$ for each $A \in \mathcal{A} \setminus \mathcal{H}_\mathcal{A}$. Lemma 5.3 gives us a direct way to do this using values that we have already computed.

The computation of $m(A)$ and $m'(A)$ for $A \in \mathcal{A}_{\text{merged}}$ can be done in time $2^r \cdot \mathrm{poly}(|x|)$. Moreover, given $m(A)$ and $m'(A)$ for each $A \in \mathcal{A}_{\text{merged}}$, the remainder of the algorithm can be carried out in polynomial time. □

The result of Theorem 5.5 provides a starting point for investigating how best to algorithmically use hierarchical structure to combine arbitrary sets of evidence. By itself, the result is restricted in various ways. Below, we discuss several suggestions for how to extend Theorem 5.5 to more general and practically useful settings.

The algorithm of Theorem 5.5 can straightforwardly be adapted also to the case where one additionally have basic belief assignments whose only focal element is the complement of some A in the merged hierarchy $\mathcal{H}_{\mathcal{A}}$–as is the case for Theorem 5.2. Therefore, one might be able to compute $\mathrm{bel}_m(A)$ for some $A \in \mathcal{A}$ more efficiently by constructing a set $\mathcal{A}' \subseteq \mathcal{A}$ such that for each focal element D of a given basic belief assignment with a single focal element, either (1) $\mathcal{H}_{\mathcal{A}'}$ contains a set in $\mathcal{A}_{\text{merged}}$ that is a superset of D, or (2) D is the complement of some set in $\mathcal{H}_{\mathcal{A}'}$. For example, take $\mathcal{A} = \{\{a,b\},\{b,c\},\{a,b,c,d\},\{d,e\},\{e\}\}$, and suppose you have basic belief assignments with their only proper focal elements in $\mathcal{A}$. Then $\mathcal{H}_{\mathcal{A}} = \{\{a,b,c,d,e\}\}$. However, you can also consider the hierarchy $\mathcal{H} = \{\{a,b,c\},\{a,b,c,d\},\{e\}\}$ to use (the extended variant) of the algorithm of Theorem 5.5 to compute $\mathrm{bel}_m(A)$ for sets $A \in \mathcal{A}$, and this would be more efficient. An interesting direction for future research would be to develop (efficient) algorithms for finding a set $\mathcal{A}' \subseteq \mathcal{A}$ that enables the most efficient use of the algorithm of Theorem 5.5.

Moreover, the given notion of a merged hierarchy $\mathcal{H}_{\mathcal{A}}$ corresponding to a set $\mathcal{A}$ of focal elements may lead to merge (nearly) all sets contained in $\mathcal{A}$ in many situations. In such situations, the algorithm of Theorem 5.5 would boil down to combining all available evidence using a brute force algorithm. It would be interesting to study more refined notions of distance to a hierarchy. One such distance measure could be to count the number of steps needed to construct a hierarchy $\mathcal{H}$ from an arbitrary set $\mathcal{A}$ of focal elements. To be effective, this construction should be done by using operations on the sets in $\mathcal{A}$ that do not require to use a brute force algorithm to deal with the resulting sets.

5.2 Applying Topological Models of Evidence

In the preceding section, we observed that combining uncertain pieces of evidence using Dempster's rule of combination requires exponential time if further constraints are not applied. Now, we analyze the computational complexity of topological models of evidence, the second method that serves as a foundation of Chapter 3. In this case, combining evidence–not uncertain evidence–can be accomplished in polynomial time. This section is dedicated to proving this statement. We will begin with some definitions and two preliminary lemmas. Note that we revert to the standard notation of this thesis, that is, given a set of possible states S, $\boldsymbol{\mathcal{E}} = \{E_1, \ldots, E_\ell\}$ is a qualitative evidence set, where $\emptyset \neq E_j \subsetneq S$ for every $j \in \{1, \ldots, \ell\}$.

Definition 5.6 (Identical support relation). Let S be a finite set of possible states and $\boldsymbol{\mathcal{E}}$ a qualitative set of evidence. For all a, $b \in S$, we define the *identical support relation* on S as

$$a \sim_{\boldsymbol{\mathcal{E}}} b \text{ if and only if, for all } E \in \boldsymbol{\mathcal{E}},\ a \in E \text{ if and only if } b \in E.$$

Lemma 5.4 *The identical support relation $\sim_{\boldsymbol{\mathcal{E}}}$ is an equivalence relation.*

Proof.

For every $a, b, c \in S$, it follows from $\sim_{\boldsymbol{\mathcal{E}}}$'s definition that $a \sim_{\boldsymbol{\mathcal{E}}} a$, if $a \sim_{\boldsymbol{\mathcal{E}}} b$ then $b \sim_{\boldsymbol{\mathcal{E}}} a$, and if $a \sim_{\boldsymbol{\mathcal{E}}} b$ and $b \sim_{\boldsymbol{\mathcal{E}}} c$ then $a \sim_{\boldsymbol{\mathcal{E}}} c$. Therefore, the relation $\sim_{\boldsymbol{\mathcal{E}}}$ is reflexive, symmetric and transitive. □

We denote the quotient set of S with respect to $\sim_{\boldsymbol{\mathcal{E}}}$ as $S/\sim_{\boldsymbol{\mathcal{E}}}$. Given $a \in S$, $[a]$ denotes the equivalence class of a. Furthermore, we consider the set $\boldsymbol{\mathcal{E}}'$ of subsets of $S/\sim_{\boldsymbol{\mathcal{E}}}$ such that $E' = \{[a] : a \in E\}$ for all $E \in \boldsymbol{\mathcal{E}}$. Note that $\boldsymbol{\mathcal{E}}$ and $\boldsymbol{\mathcal{E}}'$ have the same size.

Example 5.3 (Part 1) **Quotient Set and $\boldsymbol{\mathcal{E}}'$**

Let us consider $S = \{a, a', b, c, d, d', e\}$ and $\boldsymbol{\mathcal{E}} = \Big\{\{a, a', b\}, \{b, c, d, d'\}, \{d, d', e\}\Big\}$. Then, $a \sim_{\boldsymbol{\mathcal{E}}} a'$, $d \sim_{\boldsymbol{\mathcal{E}}} d'$ and $S/\sim_{\boldsymbol{\mathcal{E}}} = \{[a], [b], [c], [d], [e]\}$, where $[a] = \{a, a'\}$, $[b] = \{b\}$, $[c] = \{c\}$, $[d] = \{d, d'\}$, $[e] = \{e\}$, and $\boldsymbol{\mathcal{E}}' = \Big\{\{[a], [b]\}, \{[b], [c], [d]\}, \{[d], [e]\}\Big\}$.

Lemma 5.5 *The union of the singleton elements of the topology generated by $\boldsymbol{\mathcal{E}}'$ is exactly the minimum dense open set in the topology.*

Proof.

To recap, given a set of possible states S and a set $\boldsymbol{\mathcal{E}}$ of subsets of S, the topology generated by $\boldsymbol{\mathcal{E}}$ is the smallest set closed under arbitrary unions and finite intersections that contains every element of $\boldsymbol{\mathcal{E}}$, S and the empty set. In addition, an element of the topology is said to be *dense in the topology* if it has a non-empty intersection with every other non-empty element of it. In Lemma 3.1, we saw that the set of all dense open sets in the generated topology has a minimum element with respect to the subset relation $\subseteq$. Let us show that the minimum dense open set in the topology generated by $\boldsymbol{\mathcal{E}}'$ is equal to the union of the singleton elements of that topology.

Let $\tau_{\mathcal{E}'}$ be the relevant topology, M its minimum dense open set and U the union of singleton elements of $\tau_{\mathcal{E}'}$. Let us see that $M \subseteq U$. Given $[a] \in M$, if $[a] \notin U$ then there is not a collection of elements of $\mathcal{E}'$ whose intersection is equal to $[a]$, i.e., $[a]$ is not an element of the topology. Then, there exist an element of the topology T such that $\{[a]\} \subset T \subseteq M$. Without loss of generality, let us assume $T = \{[a], [b]\}$ for certain $b \in S$. If there existed an element T' of the topology such that $[a] \in T'$ but $[b] \notin T'$, then $T' \cap T = \{[a]\}$ and $\{[a]\}$ would belong to $\tau_{\mathcal{E}'}$. In addition, if there existed an element T'' of the topology such that $[b] \in T''$ but $[a] \notin T''$, then $T'' \cap T = \{[b]\}$ would belong to $\tau_{\mathcal{E}'}$ and $(M \setminus T) \cup \{[b]\}$ as well—if $M \setminus T \neq \emptyset$ then it comes from intersections and unions of elements that do not contain $[a]$. We have seen that any element of $\tau_{\mathcal{E}'}$ containing $[a]$ would also contain $[b]$, so $(M \setminus T) \cup \{[b]\}$ has non-empty intersection with every element that contains $[a]$, T'', and every other element of the topology that does not contain $[a]$ or $[b]$—for our assumption that M is a dense open set in the topology. In other words, $(M \setminus T) \cup \{[b]\}$ is a dense open set in the topology. Since $(M \setminus T) \cup \{[b]\}$ has strictly less elements than M, this contradicts our assumption that M is the minimum dense open set in $\tau_{\mathcal{E}'}$. Therefore, $[a]$ and $[b]$ belong to the same elements of $\tau_{\mathcal{E}'}$ and, consequently, a and b belong to the same elements of $\mathcal{E}$—i.e., a and b belong to the same equivalence class of the relation $\sim_{\mathcal{E}}$. Then the element $T = \{[a], [b]\}$ of $\tau_{\mathcal{E}'}$ is actually $T = \{[a]\}$, so $\{[a]\}$ is a singleton contained in $\tau_{\mathcal{E}'}$ and $[a] \in U$ as desired. Conversely, $U \subseteq M$ follows immediately, since M intersects every element of the topology, every singleton element in particular. □

Next, we define the decision problem we want to solve and present an algorithm that solves it efficiently.

TOPOLOGICAL-BELIEF
Input: A set of possible states S, a qualitative evidence set $\mathcal{E} \subseteq 2^S$, and a proposition $P \subseteq S$.
Question: $B(P) = 1$? Where $B(\cdot)$ is the binary belief operator defined in Proposition 3.9.

Theorem 5.6 *Topological-Belief can be computed in polynomial time.*

Proof.

Let S, $\mathcal{E}$ and P be the input described in Topological-Belief. By definition, $B(P)$ returns 1 if and only if there exists an element D of the topology generated by $\mathcal{E}$, $\tau_{\mathcal{E}}$, such that $D \subseteq P$ and $D \cap T \neq \emptyset$ for all non-empty element T of the topology $\tau_{\mathcal{E}}$. As before, we refer to the element of $\tau_{\mathcal{E}}$ that is dense in the topology and has the smallest cardinality among the elements of $\tau_{\mathcal{E}}$ dense in it as the minimum dense open set in $\tau_{\mathcal{E}}$.

$B(P)$ returns 1 if and only if the minimum dense open set M in $\boldsymbol{\tau_{\mathcal{E}}}$ is contained in P. The right-to-left direction is clear, since $M \subseteq P$ means that P contains a dense open set in the topology $\boldsymbol{\tau_{\mathcal{E}}}$, and therefore $B(P) = 1$. To prove the other direction, notice that $B(P) = 1$ implies that there exists an element $D \in \boldsymbol{\tau_{\mathcal{E}}}$ which is dense in the topology such that $D \subseteq P$. Since M is the minimum element of the set of all dense-open subsets of the generated topology with respect to the subset relation $\subseteq$, M is a subset of all dense open sets in $\boldsymbol{\tau_{\mathcal{E}}}$ and, in particular, $M \subseteq D$. Therefore, we will present an efficient algorithm to find the minimum dense open set in $\boldsymbol{\tau_{\mathcal{E}}}$ and to conclude whether $B(P) = 1$ we only need to check whether M is contained in P.

The first step is computing $S/_{\sim_{\mathcal{E}}}$ and $\boldsymbol{\mathcal{E}'}$ according to the equivalence relation $\sim_{\mathcal{E}}$ defined in Lemma 5.4. Considering n the size of S and ℓ the size of $\boldsymbol{\mathcal{E}}$, this operation requires a number of steps proportional to $n \cdot \ell$. The second step involves iterating over all possible singletons $\{[a]\}$ of $S/_{\sim_{\mathcal{E}}}$ and computing the intersection of all the elements of $\boldsymbol{\mathcal{E}'}$ that contains $\{[a]\}$. If this intersection is $\{[a]\}$, we keep it. After iterating over the whole set $S/_{\sim_{\mathcal{E}}}$, we define M' as the union of the elements we have kept. Finally, we define M as the union of the elements of all equivalence classes contained in M', and we check whether $M \subseteq P$.

The key part of proving that M is the minimum dense open set in $\boldsymbol{\tau_{\mathcal{E}}}$, is to prove that M' is the minimum dense open set in $\boldsymbol{\tau_{\mathcal{E}'}}$. To this end, let us see that by the process described in the preceding paragraph we get exactly all the singletons of $\boldsymbol{\tau_{\mathcal{E}'}}$. If $\{[a]\}$ is an element of the relevant topology, then there is a finite collection of elements of $\boldsymbol{\mathcal{E}'}$ whose intersection is $\{[a]\}$, i.e., $\{E' \in \boldsymbol{\mathcal{E}'} : [a] \in E'\}$. Therefore, if we take all the elements in $\boldsymbol{\mathcal{E}'}$ that contain $\{[a]\}$, in particular, we will take $\{E' \in \boldsymbol{\mathcal{E}'} : [a] \in E'\}$ and the intersection of our selection will be $\{[a]\}$ as well. On the other hand, if the intersection of all the elements of $\boldsymbol{\mathcal{E}'}$ that contain $\{[a]\}$ is $\{[a]\}$, then there exists a finite collection of elements whose intersection is $\{[a]\}$ and this singletons belongs to the topology. Now, applying Lemma 5.5 the result holds. Note that we are iterating over the set of possible states and the set of pieces of evidence. If we denote their sizes as n and ℓ, respectively, the number of steps required to conclude this process is proportional to $n \cdot \ell$. Hence, $B(P)$ can be determined in polynomial time. $\square$

Let us finish this section with a small example of the proposed algorithm.

Example 5.3 (Part 2) **Algorithm**

Let $S = \{a, a', b, c, d, d', e\}$ be a set of possible states and $\mathcal{E} = \Big\{\{a, a', b\}, \{b, c, d, d'\}, \{d, d', e\}\Big\}$ be a qualitative set of evidence pieces. The minimum dense open set in $\tau_{\mathcal{E}}$ is $\{b, d, d'\}$, let us see how to compute it applying the process defined in the proof of Theorem 5.6.

Step 1: $S/\sim_{\mathcal{E}} = \{[a], [b], [c], [d], [e]\}$ and $\mathcal{E}' = \Big\{\{[a], [b]\}, \{[b], [c], [d]\}, \{[d], [e]\}\Big\}$

Step 2: For $[a]$, $\bigcap\Big\{\{[a], [b]\}\Big\} = \{[a], [b]\}$ (dismissed)

For $[b]$, $\bigcap\Big\{\{[a], [b]\}, \{[b], [c], [d]\}\Big\} = \{[b]\}$ (kept)

For $[c]$, $\bigcap\Big\{\{[b], [c], [d]\}\Big\} = \{[b], [c], [d]\}$ (dismissed)

For $[d]$, $\bigcap\Big\{\{[b], [c], [d]\}, \{[d], [e]\}\Big\} = \{[d]\}$ (kept)

For $[e]$, $\bigcap\Big\{\{[d], [e]\}\Big\} = \{[d], [e]\}$ (dismissed)

Step 3: $M' = \{[b], [d]\}$

Step 4: $M = \{b, d, d'\}$

Now, for every $P \subseteq S$ such that $\{b, d, d'\} \subseteq P$, $B(P) = 1$.

5.3 Applying Multi-Layer Belief Model

After studying the computational complexity of Dempster's rule of combination and topological models of evidence, we wonder how complex the multi-layer belief model is from a computational point of view. In this section, we will provide a basic analysis of the computational complexity of computing degrees of belief using this model. In particular, we will describe in precise terms what computational problem we consider, and provide some initial computational complexity results that show that applying the multi-layer belief model does not increase the computational complexity compared to applying Dempster's rule of combination.

Throughout this section, we follow the notation introduced in Chapter 3, that is, S represents a set of possible states, $\mathcal{E}^Q = \{(E_1, p_1), \ldots, (E_\ell, p_\ell)\}$ represents a quantitative evidence set, $\tau_{\mathcal{E}}$ is the topology generated by the qualitative evidence set $\mathcal{E}$, $\mathcal{J}$ is a frame of justification, $f : 2^{\mathcal{E}} \to \tau_{\mathcal{E}}$ is an evidence allocation function, and $\delta(\cdot)$, $\delta_{\tau}(f, \cdot)$, $\delta_{\mathcal{J}}(f, \cdot)$ $\mathrm{Bel}_{\mathcal{J}}(f, \cdot)$ are the functions defined in equations (3.1), (3.3), (3.4) and Definition 3.2 respectively. The following is the description of the complexity problem we study in this section.

Degree-of-Belief

Input: A set of possible states S, a quantitative evidence set $\boldsymbol{\mathcal{E}}^Q$, a frame of justification $\boldsymbol{\mathcal{J}}$, an evidence allocation function f and a proposition $P \in 2^S$.

Output: $\mathrm{Bel}_{\boldsymbol{\mathcal{J}}}(f, P)$.

$\boldsymbol{\mathcal{J}}$ and the domain of the function $f : 2^{\boldsymbol{\mathcal{E}}} \to \boldsymbol{\tau}_{\boldsymbol{\mathcal{E}}}$ are in general of size exponential in the size of S (and the other parts of the input). Therefore, whenever we do not consider a fixed frame of justification or a fixed evidence allocation function, respectively, we assume that these functions are represented as (a suitable specification of) a polynomial-time computable function.

Upper bound

We begin with an upper bound on the computational complexity of the problem Degree-of-Belief in its most general form–that is, when the frame of justification and evidence allocation function are given as part of the input.

Proposition 5.7 *Degree-of-Belief is in #P, if the frame of justification $\boldsymbol{\mathcal{J}}$ has a polynomial-time decidable characteristic function and if the evidence allocation function f is polynomial-time computable, and both of these are given as part of the input (specified in a suitable format).*

Proof.

We will show that (a suitable variant of) Degree-of-Belief is in #P. Degree-of-Belief returns fractions $q \in \mathbb{Q}$, and the complexity class #P concerns functions that return natural numbers. Therefore, in the remainder of this proof, we specify a fraction $q \in \mathbb{Q}$ by two natural numbers $n, d \in \mathbb{N}$ such that $q = {}^n\!/\!_d$. One can straightforwardly extend results for #P functions that return natural numbers to #P functions that specify fractions in this way. We omit further details of this in this proof.

We will show that the function $\mathrm{Bel}_{\boldsymbol{\mathcal{J}}}(f, P)$ is in #P in several steps. In particular, we will show that $\delta(\boldsymbol{E})$, $\delta_{\boldsymbol{\tau}}(f, T)$ and $\delta_{\boldsymbol{\mathcal{J}}}(f, A)$ are computable in #P, using each result to establish the next result. Then, using these intermediate results, we will show that $\mathrm{Bel}_{\boldsymbol{\mathcal{J}}}(f, P)$ is in #P. In order to do this, we will use various closure properties of #P (Ogiwara and Hemachandra, 1993). These closure properties can, as mentioned above, be straightforwardly extended to #P functions that return fractions $q = {}^n\!/\!_d$ (by specifying n and d).

Take an input consisting of a set S of possible states, a quantitative evidence set $\boldsymbol{\mathcal{E}}^Q = \{(E_1, p_1), \ldots, (E_\ell, p_\ell)\} \subseteq 2^S \times (0, 1)$, a frame $\boldsymbol{\mathcal{J}}$ of justification, an evidence allocation function $f : 2^{\boldsymbol{\mathcal{E}}} \to \boldsymbol{\tau}_{\boldsymbol{\mathcal{E}}}$–where the frame of justification and the evidence allocation function are both given by suitably specified polynomial-time computed functions – and a proposition $P \in 2^S$.

Firstly, we show that $\delta(\boldsymbol{E})$ is in #P. Recall that $\delta(\boldsymbol{E}) = \prod_{E_j \in \boldsymbol{E}} p_j \prod_{E_j \notin \boldsymbol{E}} 1-p_j$. Clearly, each p_j and each $(1-p_j)$ is computable in #P, as they are given as part of the input. Then, since #P is closed under multiplication (over a polynomial number of #P functions), we get that $\delta(\boldsymbol{E})$ is in #P as well.

Next, let us turn to $\delta_{\boldsymbol{\tau}}(f, T)$. Recall that $\delta_{\boldsymbol{\tau}}(f, T) = \sum_{\boldsymbol{E}: f(\boldsymbol{E})=T} \delta(\boldsymbol{E})$ if $T \in \boldsymbol{\tau}_{\boldsymbol{\mathcal{E}}}$, and $\delta_{\boldsymbol{\tau}}(f, T) = 0$ otherwise. Consider the function δ' such that $\delta'(\boldsymbol{E}) = \delta(\boldsymbol{E})$ if $f(\boldsymbol{E}) = T$ and such that $\delta'(\boldsymbol{E}) = 0$ otherwise. Then, because f is a polynomial-time computable function and because $\delta(\boldsymbol{E})$ is in #P, we know that δ' is also in #P. Moreover, whenever $T \in \boldsymbol{\tau}_{\boldsymbol{\mathcal{E}}}$, it holds that $\delta_{\boldsymbol{\tau}}(f, T) = \sum_{\boldsymbol{E}} \delta'(\boldsymbol{E})$. Then, because f is a polynomial-time computable function, because δ' is in #P, and because #P is closed under addition (over an exponential number of #P functions), we can conclude that $\delta_{\boldsymbol{\tau}}(f, T)$ is in #P as well.

Next, consider $\delta_{\boldsymbol{\mathcal{J}}}(f, A)$. Recall that $\delta_{\boldsymbol{\mathcal{J}}}(f, A) = {}^{\delta_{\boldsymbol{\tau}}(f,A)}/_{\sum_{T \in \boldsymbol{\mathcal{J}}} \delta_{\boldsymbol{\tau}}(f,T)}$ if $A \in \boldsymbol{\mathcal{J}}$ and $\delta_{\boldsymbol{\mathcal{J}}}(f, A) = 0$ otherwise. Because #P is closed under addition (over an exponential number of #P functions), by a similar argument as we used above, because the characteristic function of $\boldsymbol{\mathcal{J}}$ is polynomial-time computable, we know that $\sum_{T \in \boldsymbol{\mathcal{J}}} \delta_{\boldsymbol{\tau}}(f, T)$ is in #P. Then, because #P is closed under division (of two #P functions), we can conclude that $\delta_{\boldsymbol{\mathcal{J}}}(f, A)$ is in #P as well.

Finally, let us look at $\mathrm{Bel}_{\boldsymbol{\mathcal{J}}}(f, P)$. Recall that $\mathrm{Bel}_{\boldsymbol{\mathcal{J}}}(f, P) = \sum_{A \subseteq P} \delta_{\boldsymbol{\mathcal{J}}}(f, A)$. Because #P is closed under addition (over an exponential number of #P functions), by a similar argument as we used above, we can conclude that $\mathrm{Bel}_{\boldsymbol{\mathcal{J}}}(f, P)$ is in #P. This concludes our proof that DEGREE-OF-BELIEF is in #P. □

Lower bound

Next, we show that the upper bound of #P-membership is matched by a #P-hardness lower bound, even for a particular case where we use a fixed frame of justification and a fixed evidence allocation function. In fact, this is the case that boils down to Dempster's rule of combination (see Proposition 3.8)–which we can use to straightforwardly establish #P-hardness.

Proposition 5.8 *DEGREE-OF-BELIEF is #P-hard, even when we require that the frame of justification is $\boldsymbol{\mathcal{J}}^{DS}$ and that the evidence allocation function is the function i as defined in Proposition 3.5.*

Proof.

> Consider the case where the frame of justification is $\mathcal{J}^{DS}$ and where the evidence allocation function is the function i as defined in Proposition 3.5. We will show that DEGREE-OF-BELIEF is #P-hard, even under these restrictions. By Proposition 3.8, we know that in this case, DEGREE-OF-BELIEF boils down to computing the belief $\mathrm{Bel}(P)$ of a proposition P based on applying Dempster's rule of combination to a given set of basic belief assignments with a single proper focal element each. This problem is #P-complete as shown by Theorem 5.1, and thus #P-hardness of DEGREE-OF-BELIEF follows directly. □

From this we can conclude that the problem in its most general form is #P-complete.

Corollary 5.3 *DEGREE-OF-BELIEF is #P-complete.*

5.4 Use Case Scenario: Medical Diagnosis

Shafer and Logan (2008)'s algorithm for efficiently combining hierarchical evidence via Dempster's rule of combination is motivated by the medical applications that can accept hierarchical evidence as Gordon and Shortliffe (2008) pointed out. Taking this as a starting point, let us consider a situation where we may not have a hierarchical evidence structure by default, but we may still have *near* hierarchical evidence and we can take advantage of the results of Section 5.1.2.

Continuing with the theme of medical diagnosis, we can think of cases of patients with an unusual medical history that does not follow the expected differential diagnosis to the letter—e.g., an apparently clear infection that is negative on every single specific test for bacteria, viruses, parasites or fungi. Another complication could be to have a medical history that requires the interaction of several teams, such as neurologists, immunologists, microbiologists and dermatologists. Unexpected test results or complex multidisciplinary symptoms may benefit from diagnostic support software that can compare the analysis results of the current patient with historical data and provide evidence of different causes with some degree of certainty. If this software is equipped with tools that execute the results presented in Sections 5.1.2 and 5.2, it can be used to efficiently compute belief in specific causes and help practitioners to efficiently test hypotheses and make early diagnoses. There are cases where these unexpected and multi-factorial symptoms are extremely severe and doctors have to find a solution against time. Especially in these cases, it is impossible to simply run all existing tests: there are time, resource and sample limitations. Therefore, focusing the search on one main hypothesis and not other can make all the difference.

Imagine a patient who comes to the emergency department with a very high fever for the last 72 hours. During the initial exploration and routine tests, the practitioner observes respiratory symptoms. Markers of inflammation are very high. However, every culture and PCR is negative, i.e. the specific tests for bacteria and viruses are negative. During the intake interview, the patient reported that he had eaten a bad snack the day before the appearance of the fever. In addition, the patient reported that he had been exposed to animals in the last few days. After a few days in hospital, the patient also presents neurological symptoms.

After running this case through our hypothetical diagnosis support software, we get the following set of possible states and basic belief assignments: The fever determines the set of possible states $S = \{\text{Bacterium}, \text{Virus}, \text{Autoimmune Disease}, \text{Parasite}, \text{Intoxication}\}$, summarized by $S = \{B, V, A, P, I\}$. Respiratory symptoms and positive tests provide the basic belief assignment

$$m_1(X) = \begin{cases} 0.9 & \text{if } X = \{B, V\}, \\ 0.1 & \text{if } X = S, \\ 0 & \text{otherwise.} \end{cases}$$

Markers of inflammation provide the basic belief assignment

$$m_2(X) = \begin{cases} 0.95 & \text{if } X = \{B, V, A, P\}, \\ 0.05 & \text{if } X = S, \\ 0 & \text{otherwise.} \end{cases}$$

Patient's information about ingesting a spoiled food provides

$$m_3(X) = \begin{cases} 0.2 & \text{if } X = \{I\}, \\ 0.8 & \text{if } X = S, \\ 0 & \text{otherwise.} \end{cases}$$

Patient's information about exposure to animals provides

$$m_4(X) = \begin{cases} 0.1 & \text{if } X = \{P\}, \\ 0.9 & \text{if } X = S, \\ 0 & \text{otherwise.} \end{cases}$$

Finally, the results of neurological tests provide

$$m_5(X) = \begin{cases} 0.85 & \text{if } X = \{V, A\}, \\ 0.1 & \text{if } X = S, \\ 0 & \text{otherwise.} \end{cases}$$

Our support software will reject the Shafer and Logan (2008)'s algorithm because Theorem 5.3 concludes that the available body of evidence does not form a hierarchy. We can see this by applying Theorem 5.3.(b): In our example, $\boldsymbol{\mathcal{A}} = \Big\{(\{B,V,A,P\},\{I\}), (\{B,V\},\{A,P,I\}), (\{V,A\},\{B,P,I\}), (\{P\},\{B,V,A,I\})\Big\}$ and $(\{B,V\},\{A,P,I\}) \precneq 4 \succneq (\{V,A\},\{B,P,I\})$, so Theorem 5.3 does not hold.

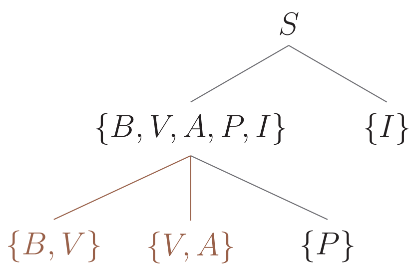


Figure 5.3: Diagram representation of the focal elements involved in the example. The colored sets are the conflicting ones.

Next, our support system may try to use Proposition 5.4 to extract the largest hierarchies from the available evidence. In this case, we can obtain a hierarchy by simply removing one element. We can get the hierarchy shown in Figure 5.4.(a) or the hierarchy shown in Figure 5.4.(b).

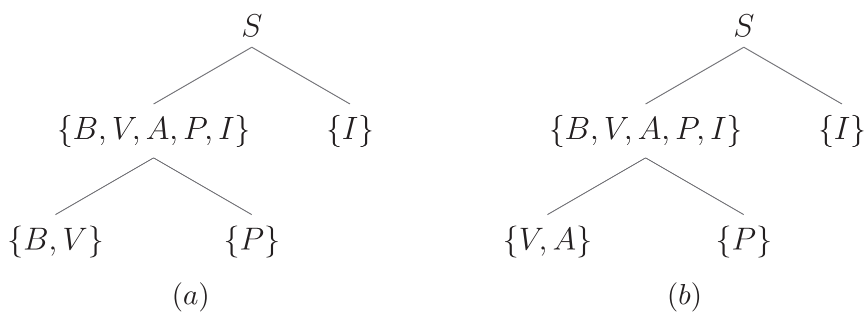


Figure 5.4: Diagram representation of the largest hierarchies within the available focal elements.

However, in both cases we remove very certain evidence that may be the key to solving the problem. Therefore, the use of Theorem 5.5 may be an interesting alternative, as it computes the exact value of the degree of belief based on Dempster's rule of combination by reducing the computation time compared to the straightforward algorithm. For this example, the merged hierarchy is the one shown in Figure 5.5 with level of merging 3. The algorithm described in the proof of Theorem 5.5 would merge two basic belief assignments by brute force and that combined belief function with the remaining ones by applying Shafer and Logan (2008)'s algorithm.

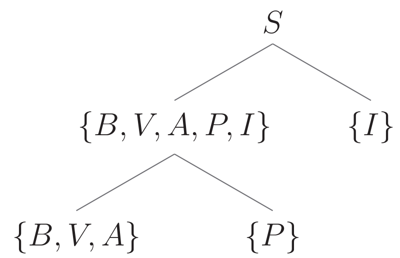


Figure 5.5: Diagram representation of the merged hierarchy obtained by applying Theorem 5.5 to the example.

Needless to say, this scenario is a toy example intended to show how alternative computations of generally computationally costly methods, such as applying Dempster's rule of combination, can be useful. This is particularly true when the conceptual and material complexity of the problem makes the additional research and engineering they require worthwhile (compared to less expressive but more efficient methods). As we have said throughout this chapter, our results are initial steps in a promising line of research, but they themselves have room for improvement. For instance, knowing the degree of belief of the sets collected in Figure 5.5 does not give too valuable information to the doctors: the strongest belief among the most specific leaves of the diagram is for $\{B, V, A\}$, which only excludes pieces of evidence with very low certainty from the search.

Another result of this chapter that may apply to this scenario is Theorem 5.6, i.e., the application of the algorithm described for Topological-Belief. In this particular case, the algorithm does not allow uncertainty values, so we will only consider the evidence that comes from positive tests. That is, $\boldsymbol{\mathcal{E}} = \{\{B, V, A, P\}, \{B, V\}, \{V, A\}\}$. The above mentioned algorithm would return {Virus} as the strongest proposition with higher belief and consistent with all available evidence. This answer may suggest to the doctors that a virus could be the cause of the illness and therefore deepen the search in that direction. Again, in most cases, doctors will not need this external help to reach these conclusions, but in very complex cases with a lot of (and not always consistent) pieces of evidence, this kind of solution may be relevant to refine the search. Some recent papers on the application of Dempster-Shafer theory to medical diagnosis are (Atitallah et al., 2022; Herawati and Eviyanti, 2021; Kusuma and Nas, 2023).

Chapter 6

Knowledge Compilation for Combining Uncertain Evidence

Up to this point, we have studied the problem of combining uncertain evidence from different perspectives. We have introduced a belief model (the multi-layer belief model) that combines the strengths of two established and independent models for belief based on evidence (Dempster-Shafer theory of belief functions and topological models of evidence). Additionally, we have explored how to use these belief models to define a qualitative logic for uncertain evidence and belief comparison, demonstrating once again their strong theoretical foundations. We have also delved into the computational limitations and possibilities for implementing these models in real-world scenarios. Building on this direction, the upcoming chapter is dedicated to presenting a practical computational approach to compute Dempster's rule of combination and its unnormalized version defined for the transferable belief model. We focus on these two rules for the significant body of work developed around their applications. Nevertheless, in the context of this thesis, this chapter serves as an example of the potential real-world applications of the discussed belief models, despite their computational complexity.

Our proposal in this chapter is to use weighted model counting and knowledge compilation to compute Dempster's rule of combination and the rule of combination of the transferable belief model. The general idea is to define a propositional formula φ that allows us to apply weighted model counting and obtain the outcome of the two mentioned rules. Needless to say, this operation does not improve the computational complexity of the problem, however, φ can be compiled into a particular circuit representation that enables weighted model counting to be carried out efficiently. That is, by compiling φ into this circuit once—and paying an expensive computational cost—we can efficiently compute the degree of belief for any proposition, even after the uncertainties change. The initial compilation only depends on the pieces of evidence we expect to have. That is, if we expect to receive (or not receive) evidence about the sets E_1, E_2 and E_3 with a variable degree of certainty, we only need to consider E_1, E_2 and E_3 in the initial compilation. After that, if we get the evidence pieces $(E_1, 0.6)$ and $(E_2, 0.3)$, both Dempster's rule of combination and the combination rule of the transferable belief model can be computed in polynomial time. In contrast, if we get an unexpected piece of evidence $(E_4, 0.7)$ we would need to start the process all over again, including the compilation step.

The chapter is organized as follows. Section 6.1 is dedicated to introducing some preliminaries necessary to follow the rest of the content. In Section 6.2, we introduce the formula φ and the weights necessary for implementing the combination rules through weighted model counting. Following this, in Section 6.3, we define the knowledge compilation approach that will permit the efficient execution of weighted model counting. Moving forward, Section 6.4 shows how applying knowledge compilation ensures polynomial time computations, including the compilation step, for hierarchical evidence. Finally, in Section 6.5, we conclude the chapter with an example illustrating how this approach could enhance overall performance compared to straightforward application of the mentioned combination rules.

6.1 Preliminaries

Given the focus of this chapter, we will adhere to Dempster-Shafer notation. Consequently, for a finite set of possible states S, we will represent the available evidence as a collection of basic belief assignments $m_1, \ldots, m_\ell$. As we saw in Chapter 2, the transferable belief model permits these mass functions to assign positive values to the empty set. For clarity, we will refer to these as *mass functions* instead of basic belief assignments when appropriate. Additionally, we will refer to Dempster's rule of combination and its unnormalized version by *normalized rule of combination* and the *unnormalized rule of combination*, respectively, in order to clarify their relationship. We recap their respective definitions in equations (6.1) and (6.2).

$$(m_1 \oplus m_2)(A) = \begin{cases} 0 & \text{if } A = \emptyset, \\ \frac{1}{K} \sum\limits_{B \cap C = A} m_1(B) \cdot m_2(C) & \text{otherwise} \end{cases} \tag{6.1}$$

where $K = \sum_{B \cap C \neq \emptyset} m_1(B) \cdot m_2(C)$.

$$(m_1 \boxplus m_2)(A) = \sum_{B \cap C = A} m_1(B) \cdot m_2(C) \tag{6.2}$$

6.1.1 Weighted Model Counting

In order to introduce weighted model counting, we briefly recap the main notions for propositional logic, and the notation that we will use. Take a countably infinite set of propositional variables. We will typically use x, y, z (possibly with indices) to denote propositional variables. Propositional formulas φ are built using propositional variables, the constants for tautology ($\top$) and contradiction ($\bot$), and the logical connectives $\neg$, $\wedge$, $\vee$, $\rightarrow$, and $\leftrightarrow$. Literals are propositional variables x or their negation $\neg x$.

The semantics of propositional logic are given by truth assignments. Let φ be a propositional formula. Then $\mathrm{Var}(\varphi)$ denotes the set of propositional variables that occur in φ. A truth assignment $\alpha : \mathrm{Var}(\varphi) \rightarrow \{0, 1\}$ assigns to each propositional variable occurring in φ a truth value—either 0 (for false), or 1 (for true). Truth of a propositional formula φ under a truth assignment α is defined in the usual way, based on the (classical) truth tables of the logical connectives. When α makes φ true, we write $\alpha \models \varphi$ and we will say that the truth assignment α satisfies φ. We will often focus on propositional logic formulas in conjunctive normal form (CNF)—that is, formulas that consist of a conjunction of disjunctions of literals. Note that every propositional formula is logically and provably equivalent to a formula in CNF (Ebbinghaus et al., 2018).

Assuming the previous notation, we define *weighted model counting* as a particular computation that takes as input a weight function over literals and a representation of a Boolean function, and computes the weighted sum of all truth assignments satisfying the Boolean function. Let φ be a propositional formula (or another representation of a Boolean function) over a set X of propositional variables. Let $\mathrm{Lit}(\cdot)$ be a function over propositional variables that returns their literals and $w : \mathrm{Lit}(X) \rightarrow \mathbb{R}$ be a weight function that assigns to each literal over X a value. Then weighted model counting for φ and w refers to computing the following number $\mathrm{WMC}(\varphi, w)$:

$$\mathrm{WMC}(\varphi, w) = \sum_{\substack{\alpha : X \rightarrow \{0,1\} \\ \alpha \models \varphi}} \prod_{\substack{\ell \in \mathrm{Lit}(X) \\ \alpha(l) = 1}} w(\ell). \tag{6.3}$$

6.1.2 Knowledge Compilation

The main idea of the field of *knowledge compilation* is to study different representation languages—e.g., for Boolean functions—and how a one-time computationally expensive translation (called *compilation*) from one language to another can be offset by repeated gains in efficiency when performing various queries or operations on the compiled representation (as opposed to performing these queries or transformations on the original representation). In order to investigate this for the class of Boolean functions, various representation languages have been investigated. A query refers to asking a particular question about the Boolean function—e.g., to ask whether the function is satisfiable—, and a transformation refers to taking one or more Boolean functions and transforming them into another function—e.g., taking the conjunction of two functions. Useful aspects to study for these representation languages, for the purposes of knowledge compilation, include (1) the computational complexity of queries and transformations, and (2) the relative succinctness of the various languages. For an overview of basic results about knowledge compilation for Boolean functions, we refer to the seminal paper by Darwiche and Marquis (2002). All terms introduced in the remaining of this section can be also found in this source.

Many representation languages used in knowledge compilation are based on Boolean circuits. A Boolean circuit consists of a direct acyclic graph with a single node that has out-degree 0 (the output node), where the leaves are annotated with propositional variables, and the internal nodes are annotated with a Boolean connective (such that the in-degree of the node is appropriate for the connective). Truth of a Boolean circuit under a truth assignment is defined by recursively assigning each node with a truth value. The leaves are assigned to the truth value specified by the truth assignment, and the truth value of the internal nodes is determined by the Boolean connective and the truth values assigned to the children of the node. The truth of the circuit is then the truth value of the output node. A Boolean circuit is in Negation Normal Form (NNF) if all negation nodes have a leaf node as child.

A class of Boolean circuits that we will refer to in this paper is the class of DNNF circuits. A DNNF circuit is a Boolean circuit in NNF that has the property of decomposability. A circuit is decomposable if for each conjunction node N it holds that for each two distinct child-nodes N_1, N_2 of N, the variables appearing in (the subcircuits under) N_1 and N_2 are disjoint. For DNNF circuits, weighted model counting can be computed in polynomial time (Kimmig et al., 2017).

An operation on representations of Boolean functions that is of interest in the area of knowledge compilation, and in particular for results in this chapter, is that of *conditioning*. Let φ be a representation of a Boolean function f on some variables X, and let α be a truth assignment to some subset $Y \subseteq X$ of variables. Then the result of conditioning φ on α is the representation of another Boolean function g on the variables $X \setminus Y$, such that for each truth assignment β to the variables in $X \setminus Y$ it holds that β satisfies g if and only if α and β combined satisfy f. Intuitively, conditioning φ on α corresponds to the Boolean function obtained from φ by assigning truth values according to α. We say that a class of

Boolean circuits supports conditioning (in polynomial time) if given a circuit N in the class and a partial truth assignment α, we can compute, in polynomial time, a circuit M that is the result of conditioning N on α. For example, the class of DNNF circuits satisfies conditioning (Darwiche and Marquis, 2002).

6.2 Combining Uncertain Evidence with Weighted Model Counting

In this section, we show how to implement the normalized and unnormalized rules of combination (equations (6.1) and (6.2), respectively) using weighted model counting. First, we define a propositional formula φ and a weight function w on the literals for the variables of φ. Then, we establish a one-to-one correspondence between the truth assignments that satisfy φ—after substituting some variables by propositional constants—and the focal elements involved in Equation (6.2). This correspondence enables us to define an alternative method for computing the combined mass function obtained by the unnormalized rule of combination, which can be extended to the normalized case. In Chapter 2, we discussed how to compute belief, plausibility, and commonality functions from a combined mass function. Alternatively, this method allows us to compute belief and commonality functions directly, without the need to compute the combined mass function as an intermediate step.

Premise 6.1 Let $S = \{s_1, \ldots, s_n\}$ be a set of possible states, C be a subset of S and $m_1, \ldots, m_\ell$ be a collection of mass functions with a single proper focal element $A_j \subset S$ for every $j \in \{1, \ldots, \ell\}$. Note that these mass functions have a maximum of two focal elements: A_j and S.

Now, let us consider a set of variables $\mathcal{V} = \{y_j | j \in \{1, \ldots, \ell\}\} \cup \{z_i | i \in \{1, \ldots, n\}\} \cup \{x_{j,i} | j \in \{1, \ldots, \ell\}, i \in \{1, \ldots, n\}\}$ and the following propositional formula:

$$\varphi = \bigwedge_{j \in \{1,\ldots,\ell\}} \Bigg(\bigg(y_j \to \Big(\bigwedge_{s_i \in A_j} x_{j,i} \wedge \bigwedge_{s_i \notin A_j} \neg x_{j,i} \Big) \bigg) \tag{6.4}$$

$$\wedge \bigg(\neg y_j \to \bigwedge_{s_i \in S} x_{j,i} \bigg) \Bigg) \tag{6.5}$$

$$\wedge \bigwedge_{i \in \{1,\ldots,n\}} \Bigg(\bigg(z_i \to \bigwedge_j x_{j,i} \bigg) \tag{6.6}$$

$$\wedge \bigg(\neg z_i \to \bigvee_j \neg x_{j,i} \bigg) \Bigg) \tag{6.7}$$

This formula φ has one variable z_i for each element of S, one variable y_j for every proper focal element of the available mass functions, and one variable $x_{j,i}$ for each pair (A, s) where A is the proper focal element of one of the available mass functions and $s \in S$.

Finally, let us define a weight function $w : \mathcal{L} \to [0, 1]$ such that for every $l \in \mathcal{L}$ literal of a variable in $\mathcal{V}$:

$$w(l) = \begin{cases} m_j(A_j) & \text{if } l = y_j \text{ for some } j \in \{1, \ldots, \ell\}, \\ 1 - m_j(A_j) & \text{if } l = \neg y_j \text{ for some } j \in \{1, \ldots, \ell\}, \\ 1 & \text{otherwise.} \end{cases} \tag{6.8}$$

Example 6.1 (Part 1) Formula φ and Weight Function

Assume $S = \{a, b, c\}$ and let m_1 and m_2 be defined as:

$$m_1(A) = \begin{cases} 0.7 & \text{if } A = \{a, b\}, \\ 0.3 & \text{if } A = S, \\ 0 & \text{otherwise.} \end{cases}$$

$$m_2(B) = \begin{cases} 0.45 & \text{if } B = \{b, c\}, \\ 0.55 & \text{if } B = S, \\ 0 & \text{otherwise.} \end{cases}$$

Then, $\mathcal{V} = \{y_1, y_2, z_1, z_2, z_3, x_{1,1}, x_{1,2}, x_{1,3}, x_{2,1}, x_{2,2}, x_{2,3}\}$, where y_1 and y_2 correspond to m_1 and m_2 respectively, and z_1, z_2 and z_3 correspond to a, b, c respectively as well. In this context,

$$\begin{aligned} \varphi = {} & (y_1 \to x_{1,1} \land x_{1,2} \land \neg x_{1,3}) \land (\neg y_1 \to x_{1,1} \land x_{1,2} \land x_{1,3}) \\ & \land (y_2 \to \neg x_{2,1} \land x_{2,2} \land x_{2,3}) \land (\neg y_2 \to x_{2,1} \land x_{2,2} \land x_{2,3}) \\ & \land (z_1 \to x_{1,1} \land x_{2,1}) \land (\neg z_1 \to \neg x_{1,1} \lor \neg x_{2,1}) \\ & \land (z_2 \to x_{1,2} \land x_{2,2}) \land (\neg z_2 \to \neg x_{1,2} \lor \neg x_{2,2}) \\ & \land (z_3 \to x_{1,3} \land x_{2,3}) \land (\neg z_3 \to \neg x_{1,3} \lor \neg x_{2,3}) \end{aligned}$$

The corresponding weight function would be:

$$w(l) = \begin{cases} 0.7 & \text{if } l = y_1, \\ 0.3 & \text{if } l = \neg y_1, \\ 0.45 & \text{if } l = y_2, \\ 0.55 & \text{if } l = \neg y_2, \\ 1 & \text{otherwise.} \end{cases}$$

Considering the weight function w and some restrictions on the truth assignments for φ, $WMC(\varphi, w)$ returns the same combination of $m_1, \ldots, m_\ell$ as equations (6.2) and (6.1), as well as the belief and commonality functions associated to them. The remainder of this section consists of a list of results that prove the previous statement.

Proposition 6.1 *Assuming Premise 6.1, let us consider:*

- $\mathcal{F}_C$ *to be the collection of* ℓ*-tuples* $(F_1, \ldots, F_\ell)$ *such that, for every* $1 \leq j \leq \ell$, F_j *is a focal element of* m_j *and* $\bigcap F_j = C$*;*
- $\varphi_{C\text{-MASS}}$ *to be the formula* φ *when* z_i *is substituted by* $\top$ *if* $s_i \in C$ *for every* $i \in \{1, \ldots, n\}$ *and by* $\bot$ *otherwise;*
- $\mathcal{A}_C$ *to be the collection of truth assignments that satisfy* $\varphi_{C\text{-MASS}}$*.*

Then, there is a one-to-one correspondence between $\mathcal{F}_C$ *and* $\mathcal{A}_C$*.*

Proof.

Let us define a function $g : \mathcal{F}_C \to \mathcal{A}_C$ such that $g\big((F_1, \ldots, F_\ell)\big) = \alpha : \mathrm{Var}(\varphi_{C\text{-MASS}}) \to \{0, 1\}$ and

$$\begin{cases} \alpha(y_j) = 1 & \text{if and only if } F_j \neq S, \\ \alpha(x_{j,i}) = 1 & \text{if and only if } s_i \in F_j. \end{cases}$$

By an abuse of notation, we denote $\alpha(x) = 1$ as $x = 1$ when $g\big((F_1, \ldots, F_\ell)\big)$ is fixed.

First, let us see the function g gives a truth assignment. Given a tuple $(F_1, \ldots, F_\ell) \in \mathcal{F}_C$, its image by g is a truth assignment for $\varphi_{C\text{-MASS}}$ since it maps a single value in $\{0, 1\}$ to every variable in $\varphi_{C\text{-MASS}}$. In addition, $g\big((F_1, \ldots, F_\ell)\big) \vDash \varphi_{C\text{-MASS}}$. The formula $\varphi_{C\text{-MASS}}$ is a conjunction of four types of clauses. The clauses of the first type (Equation (6.4)), are false if and only if $y_j = 1$ and there exists $x_{j,i} = 0$ with $s_i \in A_j$ or $x_{j,i} = 1$ with $s_i \notin A_j$. Fixing an arbitrary $j \in \{1, \ldots, \ell\}$, $y_j = 1$ implies $F_j \neq S$ for g's definition. For Premise 6.1, if m_j has a single focal element different from S, then $F_j = A_j$. Hence, the definition of g rephrases as $x_{j,i} = 0$ if and only if $s_i \notin A_j$. Since this holds for every j, $g\big((F_1, \ldots, F_\ell)\big)$ satisfies all clauses of this type. The clauses of the second type (Equation (6.5)) are false if and only if $y_j = 0$ and there exists $x_{j,i} = 0$ for that j. But according to the definition of g, $y_j = 0$ implies $x_{j,i} = 1$ for every $i \in \{1, \ldots, n\}$ since every s_i is an element of S. So $g\big((F_1, \ldots, F_\ell)\big)$ also satisfies these clauses. The clauses of the third type of clauses (Equation (6.6)), adjusted to $\varphi_{C\text{-MASS}}$, are false if and only if $x_{j,i} = 0$ for some $j \in \{1, \ldots, \ell\}$ and some $i \in \{1, \ldots, n\}$ such that $s_i \in C$. However, if $s_i \in C$, then $s_i \in F_j$ for every $j \in \{1, \ldots, \ell\}$ since $\bigcap F_j = C$ by definition of $\mathcal{F}_C$. Following the definition of g, this implies that $x_{j,i} = 1$ for every $j \in \{1, \ldots, \ell\}$. Since we are considering an arbitrary index i, this is true for every $i \in \{1, \ldots, n\}$. Finally, a clause of the last type (Equation (6.7)), adjusted to $\varphi_{C\text{-MASS}}$, is false if and only if there exists $i \in \{1, \ldots, n\}$ such that $s_i \notin C$ and $x_{j,i} = 1$ for every $j \in \{1, \ldots, \ell\}$. Since $\bigcap F_j = C$, if $s_i \notin C$ then there is a $j \in \{1, \ldots, \ell\}$ such that $s_i \notin F_j$. So $x_{j,i} = 0$ according to the definition of g. Therefore, for every $i \in \{1, \ldots, n\}$ such that $s_i \notin C$ there exists at least one $j \in \{1, \ldots, \ell\}$ such that $x_{j,i} = 0$, and $g\big((F_1, \ldots, F_\ell)\big)$ satisfies the clause. In summary, $g\big((F_1, \ldots, F_\ell)\big)$ is a truth assignment that satisfies $\varphi_{C\text{-MASS}}$ for every $(F_1, \ldots, F_\ell) \in \mathcal{F}_C$.

Now, let us prove that the function g is bijective. Let us assume that there are two ℓ-tuples $(F_1, \ldots, F_\ell)$ and $(G_1, \ldots, G_\ell)$ such that $g\big((F_1, \ldots, F_\ell)\big) = g\big((G_1, \ldots, G_\ell)\big)$. Fixing $j \in \{1, \ldots, \ell\}$, if $y_j = 0$ then $F_j = G_j = S$. Otherwise, $F_j = G_j = A_j$ since m_j has a unique proper focal element. Therefore, $(F_1, \ldots, F_\ell) = (G_1, \ldots, G_\ell)$ and g is an injective function. To show that g is surjective, let us consider α to be a truth assignment for $\varphi_{C\text{-MASS}}$ such that $\alpha \in \mathcal{A}_C$. Let us define an ℓ-tuple $(F_1, \ldots, F_\ell)$ such that $F_j = A_j$ if $\alpha(y_j) = 1$ and $F_j = S$ otherwise. Then, $(F_1, \ldots, F_\ell) \in \mathcal{F}_C$ and $g\big(F_1, \ldots, F_\ell)\big) = \alpha$ as we will prove next. To see that $(F_1, \ldots, F_\ell) \in \mathcal{F}_C$ it is enough to notice that given $s_i \in \bigcap F_j$, if $s_i \notin C$ then there exists a j such that $x_{j,i} = 0$, what leads to contradiction with clauses (6.4) or (6.5). So $\bigcap F_j \subseteq C$. In addition, if $s_i \in C$, then for definition of F_j, m_j, and clauses (6.4) and (6.6), $s_i \in F_j$ for every $j \in \{1, \ldots, \ell\}$. Therefore, $C = \bigcap F_j$ and $(F_1, \ldots, F_\ell) \in \mathcal{F}_C$. By definition of g, $g\big(F_1, \ldots, F_\ell)\big)$ is exactly α.

The previous paragraphs show that the function g is well-defined and bijective, demonstrating that there exists a one-to-one correspondence between the collections of focal elements whose intersection is C and the collections of truth assignments that satisfy $\varphi_{C\text{-MASS}}$. □

Proposition 6.2 *Assuming Premise 6.1, let us consider:*

- $\mathcal{F}_C$ *to be the collection of* ℓ*-tuples* $(F1, \ldots, F_\ell)$ *such that* F_j *is a focal element of* m_j *for every* $j \in \{1, \ldots, \ell\}$ *and* $\bigcap F_j \subseteq C$*;*
- $\varphi_{C\text{-BEL}}$ *to be the formula* φ *when* z_i *is substituted by* $\perp$ *if* $s_i \notin C$ *for every* $i \in \{1, \ldots, n\}$*;*
- $\mathcal{A}_C$ *to be the collection of truth assignments that satisfy* $\varphi_{C\text{-BEL}}$*.*

Then, there is a one-to-one correspondence between $\mathcal{F}_C$ *and* $\mathcal{A}_C$*.*

Proof.

Let us define a function $g : \mathcal{F}_C \to \mathcal{A}_C$ such that $g\Big((F_1, \ldots, F_\ell)\Big) = \alpha : \mathrm{Var}(\varphi_{C\text{-BEL}}) \to \{0, 1\}$ and

$$\begin{cases} \alpha(z_i) = 1 & \text{if and only if } s_i \in \bigcap F_j, \\ \alpha(y_j) = 1 & \text{if and only if } F_j \neq S, \\ \alpha(x_{j,i}) = 1 & \text{if and only if } s_i \in F_j. \end{cases}$$

As before, we denote $\alpha(x) = 1$ as $x = 1$ when $g\Big((F_1, \ldots, F_\ell)\Big)$ is fixed.

Note that $\mathcal{F}_C$ is now larger than in Proposition 6.1 because each $(F_1, \ldots, F_\ell) \in \mathcal{F}_C$ is required to be a subset of C, rather than equal to it. Similarly, the current set $\mathcal{A}_C$ contains the set of truth assignments from Proposition 6.1. Now, we only fix the variables z_i linked to $s_i \notin C$, whereas previously, we fixed all the variables z_i. This also explains why we define the image of variables z_i when defining g.

To prove that g gives a truth assignment in this case, we can follow the same reasoning that in the proof of Proposition 6.1 except for the clauses of the third (Equation (6.6)) and the forth (Equation (6.7)) type. The former clauses are false when for some $i \in \{1, \ldots, n\}$, $z_i = 1$ and there exists $j \in \{1, \ldots, \ell\}$ such that $x_{j,i} = 0$. However, if $z_i = 1$ then $s_i \in \bigcap F_j$. So $s_i \in F_j$ for every $F_j \in (F_1, \ldots, F_\ell)$ and $x_{j,i} = 1$ for every $j \in \{1, \ldots, \ell\}$. The later clauses, adjusted to $\varphi_{C\text{-BEL}}$, are false when for some $i \in \{1, \ldots, n\}$, $z_i = 0$ and $x_{j,i} = 1$ for every $j \in \{1, \ldots, \ell\}$ or when $s_i \notin C$ and $x_{j,i} = 1$ for every $j \in \{1, \ldots, \ell\}$. If $z_i = 0$, then $s_i \notin \bigcap F_j$. Therefore, there exist $j \in \{1, \ldots, \ell\}$ such that $s_i \notin F_j$. By definition of g, $x_{j,i} = 0$ for those indices. If $s_i \notin C$ and $x_{j,i} = 1$ for every j, then $s_i \in F_j$ for every j by definition of g. Hence $s_i \in \bigcap F_j$, but $\bigcap F_j \subseteq C$, which leads to a contradiction. In conclusion, the image of $(F_1, \ldots, F_\ell)$ under g makes the last type of clauses true, and consequently $\varphi_{C\text{-BEL}}$ as well.

A similar reasoning as in the proof of Proposition 6.1 shows that this g is bijective. Therefore, we conclude that there exists a one-to-one correspondence between the collections of focal elements whose intersection is contained in C and the collections of truth assignments that satisfy $\varphi_{C\text{-BEL}}$. □

Proposition 6.3 *Assuming Premise 6.1, let us consider:*

- $\mathcal{F}_C$ *to be the collection of* ℓ*-tuples* $(F1, \dots, F_\ell)$ *such that* F_j *is a focal element of* m_j *for every* $j \in \{1, \dots, \ell\}$ *and* $C \subseteq \bigcap F_j$*;*
- $\varphi_{C\text{-COM}}$ *to be the formula* φ *when* z_i *is substituted by* $\top$ *if* $s_i \in C$ *for every* $i \in \{1, \dots, n\}$*;*
- $\mathcal{A}_C$ *to be the collection of truth assignments that satisfy* $\varphi_{C\text{-COM}}$*.*

Then, there is a one-to-one correspondence between $\mathcal{F}_C$ *and* $\mathcal{A}_C$*.*

Proof.

This proof follows from proposing the same function g as a bijection between $\mathcal{F}_C$ and $\mathcal{A}_C$ as in Proposition 6.2. The difference between $\varphi_{C\text{-BEL}}$ and $\varphi_{C\text{-COM}}$ lies in the substitution of z_i by $\perp$ when $s_i \notin C$ in the former, whereas we substitute z_i by $\top$ when $s_i \in C$ in the later. Therefore, to prove that g gives a truth assignment in this case, we only need to check the clauses of types three (Equation (6.6)) and four (Equation (6.7)).

A clause of type three, adjusted to $\varphi_{C\text{-COM}}$, is false if and only if $z_i = 1$ and there exists a $j \in \{1, \dots, \ell\}$ such that $x_{j,i} = 0$, or if $s_i \in C$ and there exists a $j \in \{1, \dots, \ell\}$ such that $x_{j,i} = 0$. If $z_i = 1$, then $s_i \in \bigcap F_j$, so $s_i \in F_j$ for every j and $x_{j,i} = 1$ by the definition of g. If $s_i \in C$ and $x_{j,i} = 0$, then $s_i \notin \bigcap F_j$. However, $C \subseteq \bigcap F_j$ and $s_i \in C$, leading to a contradiction. Following a similar reasoning, clauses of type four are false if and only if $z_i = 0$ and $x_{j,i} = 1$ for every $j \in \{1, \dots, \ell\}$, but the definition of g implies that if $z_i = 0$, then $s_i \notin \bigcap F_j$, so there exists at least one F_j such that $s_i \notin F_j$ and $x_{j,i} = 0$.

Once again, the proof for the bijection in Proposition 6.1 can be applied to this case. We conclude that there exists a one-to-one correspondence between the collections of focal elements whose intersection contains C and the collections of truth assignments that satisfy $\varphi_{C\text{-COM}}$. □

Theorem 6.1 *Assuming Premise 6.1,* $\boxplus_{j\in\{1,\dots,\ell\}} m_j(C)$*,* $\mathrm{bel}_{\boxplus}(C)$ *and* $\mathrm{com}_{\boxplus}(C)$ *can be computed by weighted model counting.*

Proof.

Let us start by proving the result for $\boxplus_{j\in\{1,\ldots,\ell\}} m_j(C)$. Considering the weight function w defined in Equation (6.8), the propositional formula $\varphi_{C\text{-MASS}}$ defined as in Proposition 6.1 and X the set of propositional variables in $\varphi_{C\text{-MASS}}$, the weighted counting model (Equation (6.3)) establishes that

$$WMC(\varphi_{C\text{-MASS}}, w) = \sum_{\substack{\alpha:X\to\{0,1\}\\ \alpha\models\varphi_{C\text{-MASS}}}} \prod_{\substack{\ell\in\mathrm{Lit}(X)\\ \alpha(l)=1}} w(\ell).$$

For definition of our particular weight function, this translates into

$$WMC(\varphi_{C\text{-MASS}}, w) = \sum_{\substack{\alpha:X\to\{0,1\}\\ \alpha\models\varphi_{C\text{-MASS}}}} \Big(\prod_{\substack{j\in\{1,\ldots,\ell\}\\ \alpha(y_j)=1}} m_j(A_j) \prod_{\substack{j\in\{1,\ldots,\ell\}\\ \alpha(\neg y_j)=1}} (1 - m_j(y_j))\Big).$$

In Proposition 6.1 it was proved that every truth assignment α for $\varphi_{C\text{-MASS}}$ that makes it true is uniquely mapped to a tuple $(F_1, \ldots, F_\ell)$ such that each F_j is a focal element of m_j with $j \in \{1, \ldots, \ell\}$ and $\bigcap F_j = C$. Function g defined in the proof of that proposition describes the relation between the images by α and the elements of the ℓ-tuple. Attending to this, the previous formula can be written as

$$WMC(\varphi_{C\text{-MASS}}, w) = \sum_{\substack{(F_1,\ldots,F_\ell) s.t.\\ \bigcap F_j=C}} \Big(\prod_{\substack{j\in\{1,\ldots,\ell\}\\ F_j=A_j}} m_j(A_j) \prod_{\substack{j\in\{1,\ldots,\ell\}\\ F_j=S}} (1 - m_j(A_j))\Big).$$

Since every m_j has a single proper focal element, $(1 - m_j(A_j)) = m_j(S)$, so we conclude

$$WMC(\varphi_{C\text{-MASS}}, w) = \sum_{\substack{(F_1,\ldots,F_\ell) s.t.\\ \bigcap F_j=C}} \prod m_j(F_j).$$

That is, $WMC(\varphi_{C\text{-MASS}}, w) = \boxplus_{j\in\{1,\ldots,\ell\}} m_j(C)$ for every $C \subseteq S$.

With a similar reasoning we conclude that $WMC(\varphi_{C\text{-COM}}, w) = \mathrm{com}_{\boxplus}(C)$ for every $C \subseteq S$.

For belief, following a similar reasoning we conclude that

$$WMC(\varphi_{S\text{-BEL}}, w) = \mathrm{bel}_{\boxplus}(S) = 1,$$

$$WMC(\varphi_{\emptyset\text{-BEL}}, w) = \Big(m_1 \boxplus \cdots \boxplus m_\ell\Big)(\emptyset),$$

and for $C \subset S$—strictly contained—,

$$WMC(\varphi_{C\text{-BEL}}, w) = \mathrm{bel}_{\boxplus}(C) + \Big(m_1 \boxplus \cdots \boxplus m_\ell\Big)(\emptyset).$$

Therefore, for $C \subset S$, $WMC(\varphi_{C\text{-BEL}}, w) - WMC(\varphi_{\emptyset\text{-BEL}}, w) = \mathrm{bel}_{\boxplus}(C)$.

□

Example 6.1 (Part 2) Weighted Model Counting for Evidence

In the previous example we set $S = \{a, b, c\}$ and m_1 and m_2:

$$m_1(A) = \begin{cases} 0.7 & \text{if } A = \{a, b\}, \\ 0.3 & \text{if } A = S, \\ 0 & \text{otherwise.} \end{cases}$$

$$m_2(B) = \begin{cases} 0.45 & \text{if } B = \{b, c\}, \\ 0.55 & \text{if } B = S, \\ 0 & \text{otherwise.} \end{cases}$$

We saw that $\mathcal{V} = \{y_1, y_2, z_1, z_2, z_3, x_{1,1}, x_{1,2}, x_{1,3}, x_{2,1}, x_{2,2}, x_{2,3}\}$, where y_1 and y_2 correspond to m_1 and m_2 respectively, and z_1, z_2 and z_3 corresponds to a, b, c respectively as well.

Fixing $C = \{b\}$, $\varphi_{S\text{-MASS}}$ is:

$$\begin{aligned} &(y_1 \to x_{1,1} \wedge x_{1,2} \wedge \neg x_{1,3}) \wedge (\neg y_1 \to x_{1,1} \wedge x_{1,2} \wedge x_{1,3}) \\ &\wedge (y_2 \to \neg x_{2,1} \wedge x_{2,2} \wedge x_{2,3}) \wedge (\neg y_2 \to x_{2,1} \wedge x_{2,2} \wedge x_{2,3}) \\ &\wedge (\neg x_{1,1} \vee \neg x_{2,1}) \wedge (x_{1,2} \wedge x_{2,2}) \wedge (\neg x_{1,3} \vee \neg x_{2,3}) \end{aligned}$$

There exists only one truth assignment that satisfies this formula, $\alpha : \mathrm{Var}(\varphi_{S\text{-MASS}}) \to \{0, 1\}$ such that $\alpha(x) = 1$ if and only if $x \in \{y_1, y_2, x_{1,1}, x_{1,2}, x_{2,2}, x_{2,3}\}$. Therefore,

$$WMC(\varphi_{C\text{-MASS}}, w) = m_1(\{a, b\}) \cdot m_2(\{b, c\}) = (m_1 \boxplus m_2)(C).$$

Fixing $C = \{a, b\}$, $\varphi_{S\text{-BEL}}$ is:

$$
\begin{aligned}
&(y_1 \to x_{1,1} \land x_{1,2} \land \neg x_{1,3}) \land (\neg y_1 \to x_{1,1} \land x_{1,2} \land x_{1,3}) \\
&\land (y_2 \to \neg x_{2,1} \land x_{2,2} \land x_{2,3}) \land (\neg y_2 \to x_{2,1} \land x_{2,2} \land x_{2,3}) \\
&\land (z_1 \to x_{1,1} \land x_{2,1}) \land (\neg z_1 \to \neg x_{1,1} \lor \neg x_{2,1}) \\
&\land (z_2 \to x_{1,2} \land x_{2,2}) \land (\neg z_2 \to \neg x_{1,2} \lor \neg x_{2,2}) \\
&\land (\neg x_{1,3} \lor \neg x_{2,3})
\end{aligned}
$$

There exists two truth assignments that satisfies this formula, $\alpha_1 : \text{Var}(\varphi_{S\text{-BEL}}) \to \{0, 1\}$ such that $\alpha_1(x) = 1$ if and only if $x \in \{y_1, z_1, z_2, x_{1,1}, x_{1,2}, x_{2,1} x_{2,2}, x_{2,3}\}$; and $\alpha_2 : \text{Var}(\varphi_{S\text{-BEL}}) \to \{0, 1\}$ such that $\alpha_1(x) = 1$ if and only if $x \in \{y_1, y_2, z_2, x_{1,1}, x_{1,2}, x_{2,2}, x_{2,3}\}$. Therefore,

$$WMC(\varphi_{C\text{-BEL}}, w) = m_1(\{a, b\}) \cdot m_2(S) + m_1(\{a, b\}) \cdot m_2(\{b, c\}) = \text{bel}_{\boxplus}(C).$$

Finally, fixing $C = \{a\}$, $\varphi_{S\text{-COM}}$ is:

$$
\begin{aligned}
&(y_1 \to x_{1,1} \land x_{1,2} \land \neg x_{1,3}) \land (\neg y_1 \to x_{1,1} \land x_{1,2} \land x_{1,3}) \\
&\land (y_2 \to \neg x_{2,1} \land x_{2,2} \land x_{2,3}) \land (\neg y_2 \to x_{2,1} \land x_{2,2} \land x_{2,3}) \\
&\land (x_{1,1} \land x_{2,1}) \\
&\land (z_2 \to x_{1,2} \land x_{2,2}) \land (\neg z_2 \to \neg x_{1,2} \lor \neg x_{2,2}) \\
&\land (z_3 \to x_{1,3} \land x_{2,3}) \land (\neg z_3 \to \neg x_{1,3} \lor \neg x_{2,3})
\end{aligned}
$$

There exists two truth assignments that satisfies this formula, $\alpha_1 : \text{Var}(\varphi_{S\text{-COM}}) \to \{0, 1\}$ such that $\alpha_1(x) = 1$ if and only if $x \in \{y_1, z_1, z_2, x_{1,1}, x_{1,2}, x_{2,1} x_{2,2}, x_{2,3}\}$; and $\alpha_2 : \text{Var}(\varphi_{S\text{-COM}}) \to \{0, 1\}$ such that $\alpha_1(x) = 1$ if and only if $x \in \{z_1, z_2, z_3, x_{1,1}, x_{1,2}, x_{1,3}, x_{2,1}, x_{2,2}, x_{2,3}\}$. Therefore,

$$WMC(\varphi_{C\text{-COM}}, w) = m_1(\{a, b\}) \cdot m_2(S) + m_1(S) \cdot m_2(\{b, c\}) = \text{com}_{\boxplus}(C).$$

Corollary 6.1 *Assuming Premise 6.1,* $\bigoplus_{j \in \{1,\ldots,\ell\}} m_j(C)$, $\text{bel}_{\oplus}(C)$ *and* $\text{com}_{\oplus}(C)$ *can be computed by weighted model counting.*

Proof.

It is enough to notice that

$$(m_1 \oplus m_2)(A) = \begin{cases} 0 & \text{if } A = \emptyset, \\ \frac{(m_1 \boxplus m_2)(A)}{1 - (m_1 \boxplus m_2)(\emptyset)} & \text{otherwise} \end{cases}$$

and take into account Definition 2.6 to conclude that the result follows from Theorem 6.1. In particular, for $C \neq S$, $\text{bel}_{\oplus}(C) = (1 - (m_1 \boxplus m_2)(\emptyset))^{-1} \cdot \text{bel}_{\boxplus}(C)$ and, for $C \neq \emptyset$, $\text{com}_{\oplus}(C) = (1 - (m_1 \boxplus m_2)(\emptyset))^{-1} \cdot \text{com}_{\boxplus}(C)$. $\square$

6.3 Applying Knowledge Compilation

In the previous section, we have shown how merging evidence, and answering queries about the belief function based on the combined evidence, can be carried out by performing weighted model counting on a given propositional formula, with weights that reflect the uncertainty of the evidence. In this section, we will show how the framework of knowledge compilation can be employed to manage the computational cost of (the repeated use of) weighted model counting for arbitrary formulas. In particular, we will indicate what target knowledge compilation languages can be used, and under what conditions using knowledge compilation could be a promising approach.

The main reasons for why the method of knowledge compilation makes sense are captured in the following observations.

Observation 6.1 *The construction of the propositional formula φ in Section 6.2 depends only on the focal elements of the mass functions $m_1, \ldots, m_\ell$, and not on the mass values.*

Observation 6.2 *The propositional formula φ defined in Section 6.2 can be used to efficiently compute combined mass, belief and commonality numbers for given evidence pieces, assuming that the operations of conditioning and weighted model counting can be done efficiently.*

In Section 6.2, we proved that by substituting certain variables in φ with the propositional constants tautology and contradiction, we can efficiently compute the combined mass. This process is equivalent to conditioning φ on α, where $\alpha(x) = 1$ for variables substituted by tautology and $\alpha(x) = 0$ for variables substituted by contradiction.

Another important observation that helps outline in what scenarios we can sensibly use the method of knowledge compilation is the following.

Observation 6.3 *The method outlined in Section 6.2 also works for mass functions m_j where the focal element A_j gets assigned zero mass—i.e., $m_j(A_j) = 0$ and $m_j(S) = 1$. Technically, in this case A_j would not be a focal element, but by a slight abuse of terminology we say that in this case we can still associate m_j with the focal element A_j.*

Putting together these observations, we can phrase the general method of using knowledge compilation for combining uncertain evidence as follows.

Approach 6.1 *Let $\mathcal{L}$ be a (complete) representation language for Boolean functions that has the properties that* (1) *conditioning can be computed in polynomial-time for $\mathcal{L}$ and* (2) *weighted model counting for statements $L \in \mathcal{L}$ can be computed in polynomial time.*

Take a situation where we do not (yet) have concrete pieces of evidence, but where we know the set $S = \{s_1, \dots, s_n\}$ of possible states and where we know a set of focal elements $A_1, \dots, A_\ell$ such that all evidence pieces have one such A_j as focal element.

Then we can construct the formula φ as described in Section 6.2 for the focal elements $A_1, \dots, A_\ell$, and compile φ into a statement $L \in \mathcal{L}$. This compilation is a one-time, (potentially) computationally expensive operation. After compilation, we can use L to combine any set of uncertain evidence pieces over focal elements included in $A_1, \dots, A_\ell$ in polynomial time and query the mass, belief and commonality numbers (of the combined evidence) for arbitrary propositions $P \subseteq S$ in polynomial time. Here polynomial time is measured in the size of L and the size of the representation of the evidence pieces.

One representation language $\mathcal{L}$ that satisfies the required properties is that of Boolean DNNF circuits. In fact, there are off-the-shelf compilation algorithms available to compile (arbitrary) propositional formulas into DNNF circuits.

Approach 6.1 has various strengths, but also several limitations. We will briefly reflect on both the advantages and disadvantages of the approach.

The most prominent advantage of the approach is that once we have performed the compilation phase—after having decided on the set S of states and the (possible) focal elements $A_1, \dots, A_\ell$—all further operations and queries can be done efficiently. For example, we can add evidence (any set among $A_1, \dots, A_\ell$) or change the weight of previous evidence, and (re-)compute the combined mass, belief or commonality functions efficiently.

In other words, it works very well in circumstances where the set of possible evidence pieces is known beforehand, and where the actual evidence changes frequently. Think, for example, of a setting where there is a fixed set of sensors—each able to provide uncertain evidence for a fixed focal element—and where over time the values that these sensors provide change.

In particular, we can change previously used weights to zero, which efficiently allows us to forget previously considered evidence. This can be very beneficial in settings where the degree of conflict between evidence pieces gets very high, and where one wants to decrease this degree of conflict by disregarding some previously obtained pieces of evidence. Our approach allows for efficiently finding out which previously considered evidence leads to a maximum decrease in degree of conflict when forgotten.

Another advantage of our approach is that it is independent of the exact choice of target compilation language. It works with any compilation language that supports conditioning and weighted model counting in polynomial time. This choice in compilation language allows one to experiment with which one is the most effective for any given set of (possible) focal elements that are relevant in a particular application. Moreover, the bulk of the computational difficulty is dealt with by the compilation algorithms. By using off-the-shelf, optimized compilation algorithms, we can use their efficiency and engineering without reinventing the wheel. In fact, our approach will benefit from future improvements in compilation algorithms, so our approach will perform better as research on knowledge compilation advances.

Our approach also has some limitations. Firstly, and most prominently, there is no guarantee on the running time of the compilation phase—and on the size of the result of the compilation. This means that compilation could take exponential time, and that the result is of exponential size. This would constitute a significant obstacle towards practical feasibility. To put this in perspective, such hurdles are unavoidable (under widely believed complexity-theoretic assumptions), because the problem of computing belief according to the assumed rules of combination is computationally intractable—for the normalized rule, this is a #P-complete problem (Theorem 5.1).

Another important limitation is that in order for our approach to work, one has to decide on the (possible) focal elements $A_1, \dots, A_\ell$ beforehand, before one carries out the compilation phase. This means that in settings where it is unclear beforehand what types of evidence will become available later in time, our approach is not well-suited.

Lastly, our approach only works for mass functions with a single proper focal element. In principle, our approach can be adapted to non-simple support functions as well, but this would require additional research, and it might provide an additional burden on the compilation algorithms.

6.4 Hierarchical Evidence: An Efficient Case

We have shown that computing the normalized and unnormalized rules of combination is compilable to P, meaning that they can be computed in two steps: a generic off-line phase that requires exponential time, and a specific on-line phase that requires polynomial time. In particular, we can construct a DNNF circuit in exponential time based on certain potential focal elements. Subsequently, we can efficiently solve queries, such as $\mathrm{bel}_{\oplus}(C)$, in polynomial time for specific uncertain evidence.

When the type of evidence that will be collected is known from the design phase, this approach has an advantage over computing $\mathrm{bel}_{\oplus}(C)$ using equations (6.1) or (6.2), as these computations would require exponential time for each query. However, there are specific algorithms to compute these combination rules when the evidence satisfies some specific constraints (Barnett, 1981; Bergsten and Schubert, 1993; Shafer and Logan, 2008; Shafer et al., 1987). In Chapter 5 we explored the case for hierarchical evidence (Shafer and Logan, 2008). Now, we show that Approach 6.1 can be also implemented in polynomial time for this specific structure of evidence.

To prove it, we will first define a linear order $<$ among all the focal elements introduced in the input. This order will allow us to define an order among the propositional variables $\mathcal{V}$ introduced in Section 6.2, and build an ordered binary decision diagram ($\mathrm{OBDD}_<$)—which is a specific type of DNNF circuit. Next, we will propose an $\mathrm{OBDD}_<$ sentence equivalent to the propositional formula φ defined in Section 6.2. Finally, we will show that building this diagram requires a polynomial number of steps. This confirms that translating φ into such $\mathrm{OBDD}_<$ sentences and applying weighted model counting efficiently solves the problem of merging evidence using the normalized and unnormalized rules of combination.

Premise 6.2 Let $S = \{s_1, \ldots, s_n\}$ be a set of possible states, $m_1, \ldots, m_\ell$ be a collection of mass functions with a single proper focal element $A_j \subset S$ for every $j \in \{1, \ldots, \ell\}$, and $\mathcal{F}$ be the set of the proper focal elements of the previous mass functions. Let us assume that the set $\mathcal{F} \cup \{S\}$ is a hierarchy. That is, for all F_j and F_k in $\mathcal{F} \cup \{S\}$ it holds that: if $F_j \cap F_k \neq \emptyset$, then $F_j \subseteq F_k$ or $F_k \subseteq F_j$.

Example 6.2 (Part 1) Hierarchy of Focal Elements

Let $S = \{a, b, c, d\}$ be a set of possible states and m_1, m_2, m_3 and m_4 be four mass functions such that

$$\begin{array}{ll} m_1(\{a\}) = 0.6 & m_1(S) = 0.4 \\ m_2(\{c\}) = 0.4 & m_2(S) = 0.6 \\ m_3(\{a, b\}) = 0.8 & m_3(S) = 0.2 \\ m_4(\{a, b, c\}) = 0.5 & m_4(S) = 0.5 \end{array}$$

Then, the set $\mathcal{F} = \Big\{\{a\}, \{c\}, \{a, b\}, \{a, b, c\}\Big\} \cup \{S\}$ is a hierarchy.

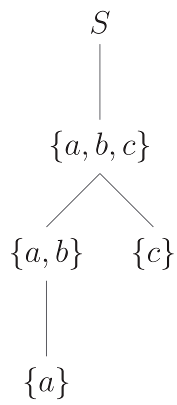


Figure 6.1: Hierarchy example.

As we mentioned previously, we will start by defining which kind of order among the proper focal elements of $\mathcal{F}$ we need.

Definition 6.1 (Valid Order $(\mathcal{F}, <)$). Assuming Premise 6.2, an order $<$ is valid over the set $\mathcal{F} \cup \{S\}$ if and only if, for every $j, k \in \{1, \ldots, \ell\}$, $F_j \subset F_k$ implies $F_j < F_k$.

Example 6.2 (Part 2) **Valid Order among Focal Elements**

Following the previous example, if we rename the elements of $\mathcal{F}$ as $A_1 = \{a\}$, $A_2 = \{a, b\}$, $A_3 = \{c\}$ and $A_4 = \{a, b, c\}$, and we order them according to their index as in Equation (6.9), then $(\mathcal{F}, <)$ is a valid order.

$$\{a\} < \{a, b\} < \{c\} < \{a, b, c\} \tag{6.9}$$

However, if we rename the elements of $\mathcal{F}$ as $A_1 = \{a\}$, $A_2 = \{a, b\}$, $A_3 = \{a, b, c\}$ and $A_4 = \{c\}$ instead, $(\mathcal{F}, <')$ is not a valid order as we can see in Equation (6.10).

$$\{a\} <' \{a, b\} <' \{a, b, c\} <' \{c\} \tag{6.10}$$

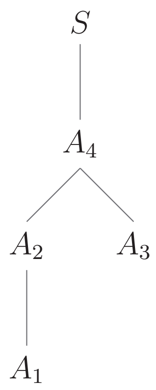


Figure 6.2: Hierarchy of focal elements in a valid order corresponding to Equation (6.9). We can see that the index of a child is always lower than the index of its parent.

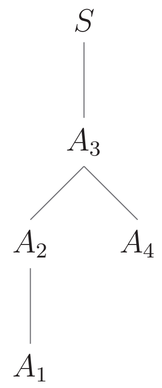


Figure 6.3: Hierarchy of focal elements in an invalid order corresponding to Equation (6.10). We can see that the index of A_4 is greater than the index of its parent.

Our goal is to create an $\mathrm{OBDD}_<$ sentence that is equivalent to the propositional formula φ defined in Section 6.2. A sentence of the compilation language $\mathrm{OBDD}_<$ is an ordered binary decision diagram. A binary decision diagram is a directed, acyclic graph with a single root node (i.e., with no incoming edges), where (i) each leaf node is labelled with 1 or 0, (ii) each non-leaf node is labelled with a propositional variable, and (iii) each non-leaf node has exactly two outgoing edges (one corresponding to 0 and one corresponding to 1). (Such a diagram can be interpreted as a representation of a Boolean function in the natural way, i.e., for each truth assignment, following the corresponding path through the graph, leading to the outcome of the function.) A binary decision diagram is said to be ordered if the propositional variables are ordered such that for every parent node, its child nodes have variables that are smaller than the variable at the parent node according to a predefined ordering. This compilation language can be seen as a proper subset of DNNF (Darwiche and Marquis, 2002), so weighted model counting can be efficiently applied to its sentences (Kimmig et al., 2017).

To achieve our goal, we define an order over the set of variables of φ, $\mathcal{V}$, based on the previous notion of a valid order. This will enable us to construct the desired $\mathrm{OBDD}_<$.

Definition 6.2 $((\mathcal{V}, <))$. Assuming Premise 6.2, let $\mathcal{V} = \{y_j | j \in \{1, \ldots, \ell\}\} \cup \{z_i | i \in \{1, \ldots, n\}\} \cup \{x_{j,i} | j \in \{1, \ldots, \ell\}, i \in \{1, \ldots, n\}\}$ be a set of propositional variables and $<$ a linear order on $\mathcal{V}$ such that:

1. $(\mathcal{F}, <)$ is a valid order, that is, $<$ orders the indexes in $\{1, \ldots, j, \ldots, \ell\}$ forming a valid order according to Definition 6.1.
2. Fixing $j \in \{1, \ldots, \ell\}$, $<$ orders the propositional variables in $\{x_{j,i} | j \in \{1, \ldots, \ell\}, i \in \{1, \ldots, n\}\}$ satisfying $s_i < s_{i'}$ for every $s_i \in A_j$ and every $s_{i'} \notin A_j$.
3. The order among propositional variables in $\{z_i | i \in \{1, \ldots, n\}\}$ is arbitrary but fixed.
4. For every $j \in \{1, \ldots, \ell\}$ and every $i \in \{1, \ldots, n\}$, $y_j < z_i$ and $x_{j,i} < z_i$.
5. For every $j, j' \in \{1, \ldots, \ell\}$ and every $i \in \{1, \ldots, n\}$, if $j < j'$ then $y_j < y_{j'}$ and $x_{j,i} < y_{j'}$.

By taking $\sigma_j : S \to S$ to be a permutation on the elements of S for each $j \in \{1, \ldots, \ell + 1\}$, the previous order would look like this:

$$
\begin{aligned}
&y_1 < x_{1,\sigma_1(1)} < \cdots < x_{1,\sigma_1(*_1)} < \cdots < x_{1,\sigma_1(n)} < \\
&y_2 < x_{2,\sigma_2(1)} < \cdots < x_{2,\sigma_2(*_2)} < \cdots < x_{2,\sigma_2(n)} < \\
&\cdots \\
&y_j < x_{j,\sigma_j(1)} < \cdots < x_{j,\sigma_j(*_j)} < \cdots < x_{j,\sigma_j(n)} < \\
&\cdots \\
&y_\ell < x_{\ell,\sigma_\ell(1)} < \cdots < x_{\ell,\sigma_\ell(*_\ell)} < \cdots < x_{\ell,\sigma_\ell(n)} < \\
&z_{\sigma_{\ell+1}(1)} < \cdots < z_{\sigma_{\ell+1}(i)} < \cdots < z_{\sigma_{\ell+1}(n)}
\end{aligned}
$$

where for each $j \in \{1, \ldots, \ell\}$, $*_j$ represents the last element $s \in S$ such that $s \in A_j$ according to the linear order $<$, i.e., if $x_{j,i} < x_{j,*_j} < x_{j,i'}$ then $s_i, s_{*_j} \in A_j$ but $s_{i'} \notin A_j$. By abuse of notation, we will avoid the sub-index of $*_j$ wherever j is clear by context. We will do so with the sub-index of σ_j.

Example 6.2 (Part 3) $(\mathcal{V}, <)$

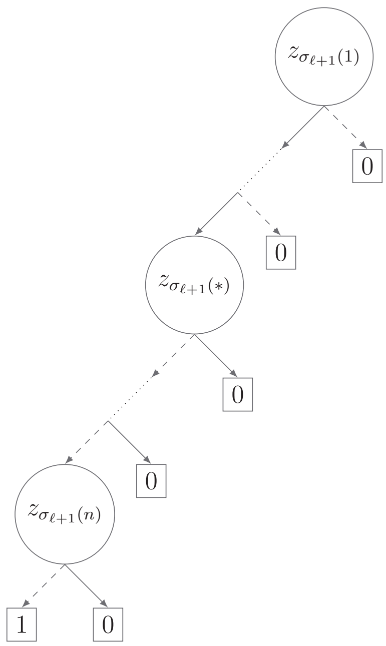
Considering $(\mathcal{F}, <)$ from the previous example and renaming the possible states as $s_1 = a$, $s_2 = b$, $s_3 = c$ and $s_4 = d$, the ordered set $(\mathcal{V}, <)$ looks like this:

$$
\begin{aligned}
&y_1 < x_{1,1} < x_{1,2} < x_{1,3} < x_{1,4} < \\
&y_2 < x_{2,1} < x_{2,2} < x_{2,3} < x_{2,4} < \\
&y_3 < x_{3,3} < x_{3,1} < x_{3,2} < x_{3,4} < \\
&y_4 < x_{4,1} < x_{4,2} < x_{4,3} < x_{4,4} < \\
&z_1 < z_2 < z_3 < z_4
\end{aligned}
$$

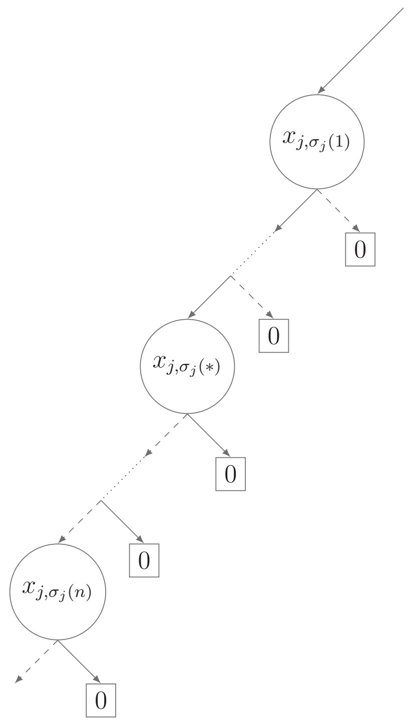
Now, we are ready to define our proposal of the $\text{OBDD}_<$ sentence equivalent to φ. To do this, we will start by defining a collection of partial branches and sub-diagrams of ordered binary decision diagrams that will be used to build our $\text{OBDD}_<$ sentence. For clarity, we will define these partial branches and sub-diagrams using visuals.

Definition 6.3 (Partial branches and sub-diagrams of $\text{OBDD}_<$). Given $(\mathcal{V}, <)$ as in Definition 6.2, we define:

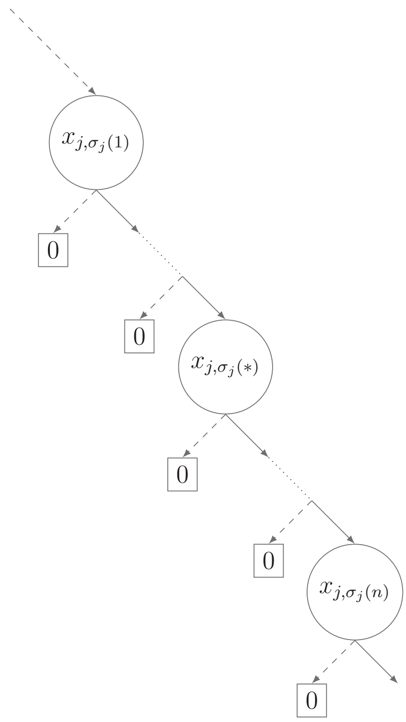
- The sub-diagram Z, where $z_{\sigma_{\ell+1}(i)} \in S$ is *true* if and only if, for all $j \in \{1, \ldots, \ell\}$, $x_{j,\sigma_{\ell+1}(i)}$ is *true* in the branch from node y_1 to node Z ($y_1 \rightarrow^* Z$):



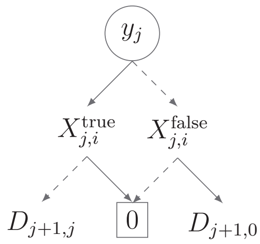
- The partial branch $X_{j,i}^{\text{true}}$. Note that $s_{\sigma_j(n)}$ is never in A_j, for every $j < \ell+1$:



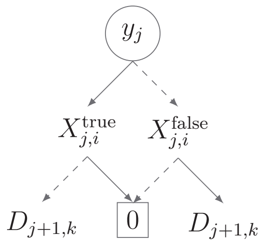
- The partial branch $X_{j,i}^{\text{false}}$. Note that false arrows lead to $\boxed{0}$ for every $x_{j,\sigma_j(i)}$ in this case:



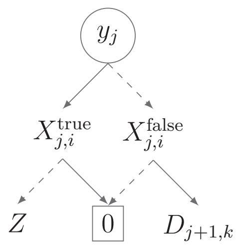
- The sub-diagram $D_{j,0}$ for $1 \leq j < \ell$ following $(\mathcal{V}, <)$ order:



- The sub-diagram $D_{j,k}$ for $1 \leq j < \ell$, $1 \leq k \leq \ell$ following $(\mathcal{V}, <)$ order, and $A_k \subseteq A_j$:



- The sub-diagram $D_{j,k}$ for $1 \leq j < \ell$, $1 \leq k \leq \ell$ following $(\mathcal{V}, <)$ order, and $A_j \cap A_k = \emptyset$. Note that in this case, Z ends on $\boxed{1}$ if and only if z_i is *false* for every $i \in \{1, \ldots, n\}$. This is because the k index is updated after $D_{k,0}$ when y_k is *true*, therefore, at this point of the diagram both y_j and y_k are *true* and $A_j \cap A_k = \emptyset$:



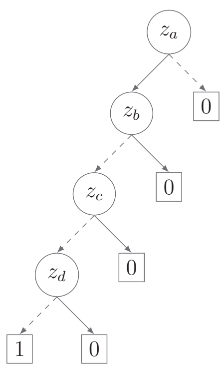
- The sub-diagram $D_{\ell+1,k}$: sub-diagram Z.
- The sub-diagram $D_{\ell+1,0}$: sub-diagram Z.

Definition 6.4 (OBDD$_<$ sentence). Assuming Premise 6.2, when applying Definition 6.3 starting by $D_{1,0}$, we obtain the OBDD$_<$ sentence that we claim is equivalent to φ.

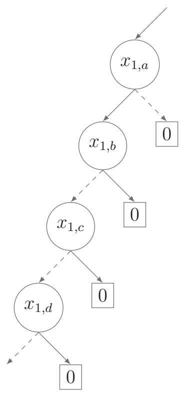
Before proving this statement, let us see how the $\mathrm{OBDD}_<$ sentence for our running example would look like.

Example 6.2 (Part 4) $\mathrm{OBDD}_<$ **sentence**

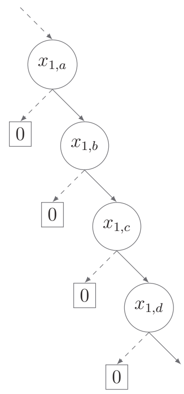
In Figure 6.4, we find the partial branches and sub-diagrams corresponding to the running example. For clarification, we specify indexes $i \in \{1, \ldots, n\}$ with the corresponding element $s_{\sigma_j(i)}$, i.e, a, b, c or d, instead of the corresponding natural numbers. The $\mathrm{OBDD}_<$ sentence of this example represented in terms of sub-diagrams is displayed in figures 6.5 and 6.6.



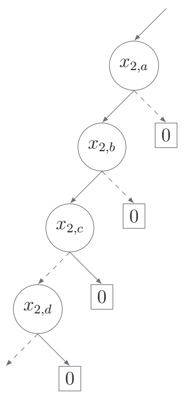
(a) Z for branch $D_{1,0} \rightarrow D_{2,1} \dashrightarrow D_{3,1} \rightarrow Z$.



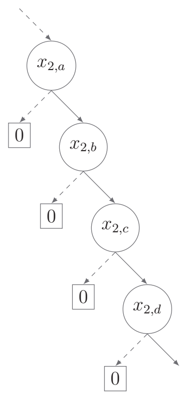
(b) $X^{\text{true}}_{1,i}$



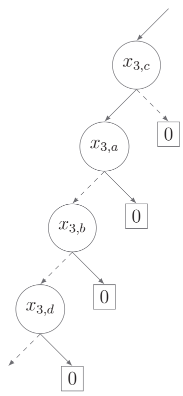
(c) $X^{\text{false}}_{1,i}$



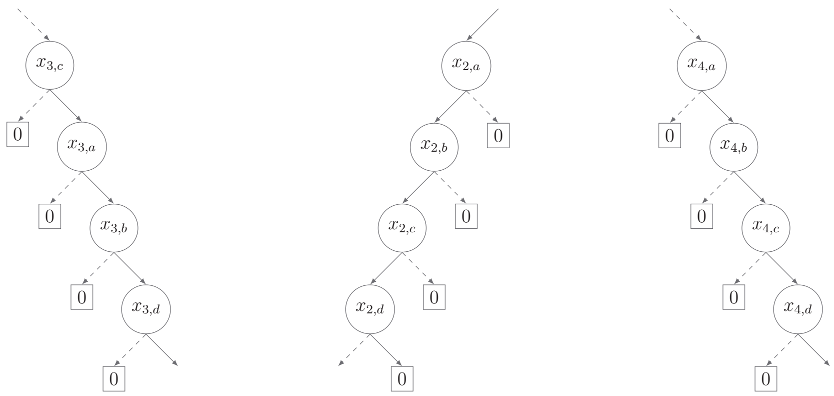
(d) $X^{\text{true}}_{2,i}$



(e) $X^{\text{false}}_{2,i}$



(f) $X^{\text{true}}_{3,i}$



(*g*) $X_{3,i}^{\text{false}}$ (*h*) $X_{4,i}^{\text{true}}$ (*i*) $X_{4,i}^{\text{false}}$

Figure 6.4: Partial branches and sub-diagrams of the OBDD$_<$ sentence of the running example.

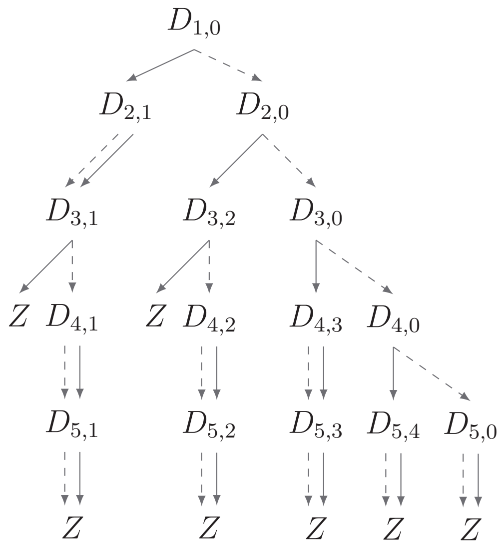


Figure 6.5: Partial branches of the OBDD$_<$ sentence of the running example.

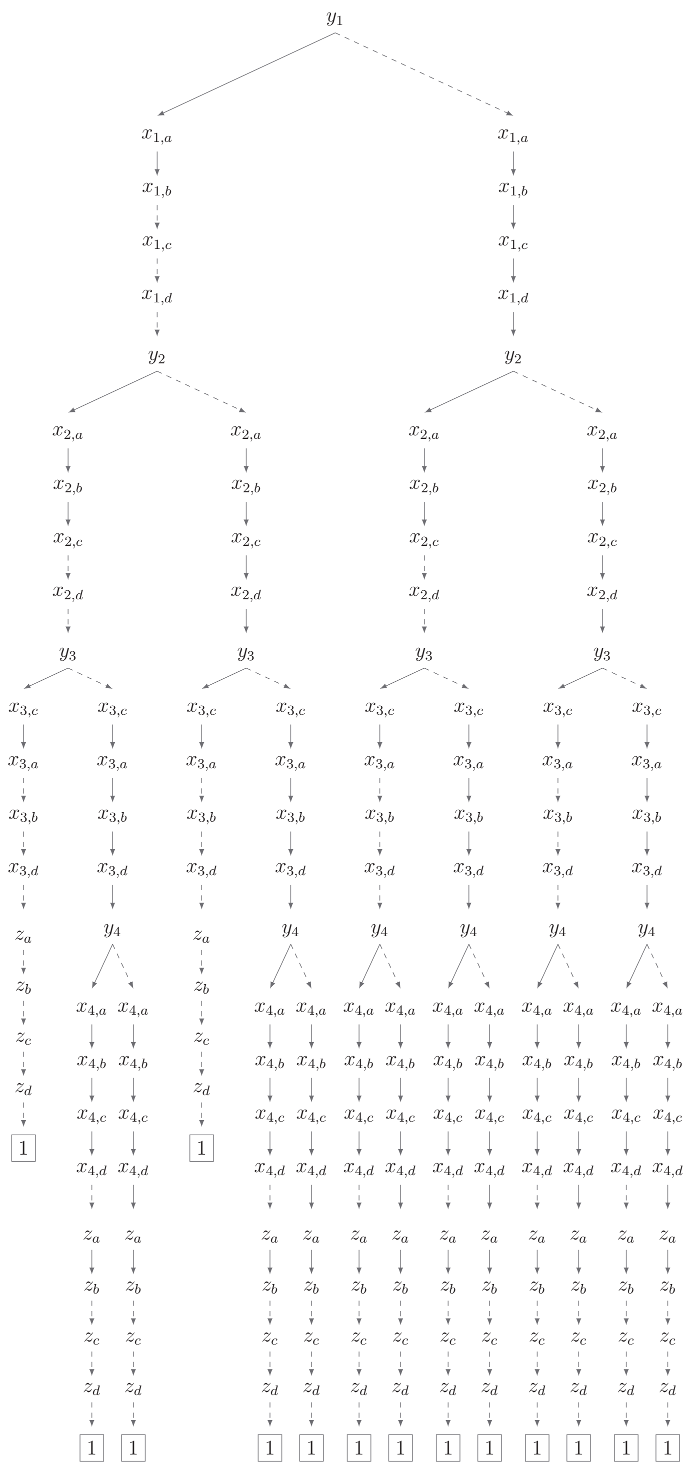


Figure 6.6: $\text{OBDD}_<$ sentence of the running example. Missing arrows lead to $\boxed{0}$.

Proposition 6.4 *Assuming Premise 6.2, the* $\mathrm{OBDD}_<$ *sentence defined above is equivalent to* φ.

Proof.

Note that given the set $\mathcal{F}$, φ and our $\mathrm{OBDD}_<$ sentence share the same propositional variables. Now, let us see that given a truth assignment α for our $\mathrm{OBDD}_<$ sentence, α satisfies φ too. According to the definition of the sub-diagram Z (see Definition 6.3), α makes z_i *true* if and only if such truth assignment makes $x_{i,j}$ *true* for every $j \in \{1, \ldots, \ell\}$ that has been seen before in the branch. There are two possible situations. Either Z occurs after seeing all y_j, or Z occurs within a $D_{j,k}$ where $A_j \cap A_k = \emptyset$. The former, ensures by definition that clauses 6.6 and 6.7 of φ are satisfied. The later, implies that both y_j and y_k are true in the branch, by definition of the corresponding sub-diagram. Since A_j and A_k do not share elements, there is at least one $x_{j,i}$ that is *false* in the branch, and therefore every z_i is *false* in it. This ensures that clauses 6.6 and 6.7 of φ are satisfied in this case as well. The definition of $X_{j,i}^{\text{true}}$ and $X_{j,i}^{\text{false}}$ (see Definition 6.3) make α to satisfy clauses 6.4 and 6.5 of φ, hence α is a truth assignment for φ.

To prove the opposite direction, we need to show that given a truth assignment α of φ, there exists a branch $y_1 \rightarrow^* Z$ that gives the same truth values to the propositional variables as α. Since every variable y_j is connected to the variable y_{j+1}, either if the former is true or false, we can build a branch following the truth values assigned by α. This branch will end in Z, and therefore in 1, since the only way to finish in 0 is contradicting α for some of the variables $x_{j,i}$ or z_i. $\square$

Corollary 6.2 *Assuming Premise 6.2 and* $C \subseteq S$, $\bigoplus_{j\in\{1,\ldots,\ell\}} m_j(C)$ *and* $\boxplus_{j\in\{1,\ldots,\ell\}} m_j(C)$ *can be computed in polynomial time.*

Proof.

Proposition 6.4 and Approach 6.1 prove that by applying weighted model counting to the $\mathrm{OBDD}_<$ sentence defined in this section, we compute $\bigoplus_{j\in\{1,\ldots,\ell\}} m_j(C)$ and $\boxplus_{j\in\{1,\ldots,\ell\}} m_j(C)$ in polynomial time. However, we need to prove that this $\mathrm{OBDD}_<$ sentence can be built in a polynomial number of steps in order to claim that we can efficiently compute $\bigoplus_{j\in\{1,\ldots,\ell\}} m_j(C)$ and $\boxplus_{j\in\{1,\ldots,\ell\}} m_j(C)$ from the input. Considering n the number of elements of S and ℓ the number of mass functions described in Premise 6.2, this $\mathrm{OBDD}_<$ sentence involves $p(\ell)$ sub-diagrams $D_{j,k}$ and each of them contains $p(n)$ nodes—where $p(x)$ represents a polynomial of variable x. Therefore, the proposed $\mathrm{OBDD}_<$ sentence can be constructed in a polynomial number of steps. $\square$

In summary, Approach 6.1 for $\mathcal{L} =$ OBDD is an alternative to Shafer and Logan's algorithm for implementing the normalized rule of combination on hierarchies. One difference with respect to their algorithm is that they consider dichotomous basic belief assignments, whereas our solution is restricted to simple basic belief assignments. In another respect, this approach extends Shafer and Logan's algorithm since it is also valid for the unnormalized rule of combination, and the query $\mathrm{bel}_{\oplus}(\cdot)$ can be implemented for any $P \subseteq S$. Shafer and Logan's algorithm, in contrast, is defined only for P included in the hierarchy.

6.5 Use Case Scenario: Air Traffic Control

The solution presented in this chapter for combining uncertain evidence by applying Dempster's rule of combination, or its unnormalized version, is particularly relevant when

1. The use case scenario has a design phase and an execution phase—e.g., programming a vacuum cleaner robot vs. using a vacuum cleaner robot at home.
2. The focal elements of the potentially collected basic belief assignments are known from the design phase—e.g., a vacuum cleaner robot is programmed to detect walls and obstacles and nothing else.
3. Each piece of evidence is associated with a degree of certainty—e.g., the vacuum cleaner robot can collect information about 'there is a wall here' with 80% certainty, even though it could be a wardrobe or something else.
4. Each source—e.g., sensors, output of deep learning algorithms, human reports, etc.—provides a piece of evidence.
5. The execution phase is performed several independent times—e.g., each time we use the vacuum cleaner robot.
6. There may be corrections to the original input that require us to forget some of the collected evidence—e.g., modifying the vacuum cleaner robot's room map to forget the evidence 'there is a carpet' because there is just a different type of floor.

Items 3 and 4 refer to modelling the problem with Dempster-Shafer theory and the constraint of only allowing simple basic belief assignments. Items 1 and 2 are specific constraints of the approach described in this chapter. Items 5 and 6 are scenarios that may particularly benefit from the proposed solution.

As the vacuum cleaner robot example suggests, we can find suitable scenarios by thinking about situations that require merging information from different sensors (and possibly other sources). For example, let us consider an air traffic situation.

Imagine a group of air traffic controllers at an international airport monitoring incoming flights. Suppose the air traffic control support system is equipped with the DNNF circuit described in this chapter for focal elements determined by the variables of interest for air traffic controllers. Assume they collect information on

(a) potential landing anomalies due to weather conditions,

(b) potential landing anomalies due to deviation from the optimum glide path,

(c) potential landing anomalies for historical data,

(d) potential landing anomalies for runway condition,

(e) potential landing anomalies for pilot warnings.

Let $S = \{\text{Severe Weather}, \text{Runway Sub-optimal Condition}, \text{Fail Flight Control Sensors}, \text{Sub-optimal Pilot Performance}\}$ be a set of possible states. We will summarize it as $S = \{W, R, F, P\}$ (W for weather, R for runway, F for flight and P for pilot). The previous information is then translated into the following potential focal elements: Item (a) provides evidence for anomalies for bad weather conditions and sub-normal runway conditions (since weather conditions affect the state of the runway). We will denote it by $\{W, R\}$. Item (b) provides evidence for anomalies due to an issue in the flight control sensors or due to the pilot performance. We will denote it as $\{F, P\}$. Let us say that item (c) shows a correlation between landing anomalies with the set $\{\text{Runway Sub-optimal Condition}, \text{Fail Flight Control Sensors}\}$, which is evidence for $\{R, F\}$. Item (d) refers to evidence coming from runway inspections and we will formalize it as evidence for $\{R\}$. Finally, item (e) refers to evidence coming from pilot reports and we will formalize it as evidence for $\{P\}$.

According to these potential focal elements, the programmers of the air traffic control support system had to consider the propositional formula

$$
\begin{aligned}
\varphi = {} & (y_{(a)} \rightarrow (x_{(a),W} \wedge x_{(a),R} \wedge \neg x_{(a),F} \wedge \neg x_{(a),P})) \\
& \wedge (\neg y_{(a)} \rightarrow (x_{(a),W} \wedge x_{(a),R} \wedge x_{(a),F}) \wedge x_{(a),P})) \\
& \wedge (y_{(b)} \rightarrow (x_{(b),F} \wedge x_{(b),P} \wedge \neg x_{(b),W} \wedge \neg x_{(b),R})) \\
& \wedge (\neg y_{(b)} \rightarrow (x_{(b),F} \wedge x_{(b),P} \wedge x_{(b),W}) \wedge x_{(b),R})) \\
& \wedge (y_{(c)} \rightarrow (x_{(c),R} \wedge x_{(c),F} \wedge \neg x_{(c),W} \wedge \neg x_{(c),P})) \\
& \wedge (\neg y_{(c)} \rightarrow (x_{(c),R} \wedge x_{(c),F} \wedge x_{(c),W}) \wedge x_{(c),P})) \\
& \wedge (y_{(d)} \rightarrow (x_{(d),R} \wedge \neg x_{(d),W} \wedge \neg x_{(d),F} \wedge \neg x_{(d),P})) \\
& \wedge (\neg y_{(d)} \rightarrow (x_{(d),R} \wedge x_{(d),R} \wedge x_{(d),W} \wedge x_{(d),F}) \wedge x_{(d),P})) \\
& \wedge (y_{(e)} \rightarrow (x_{(e),P} \wedge \neg x_{(e),W} \wedge \neg x_{(e),R}) \wedge \neg x_{(e),F})) \\
& \wedge (\neg y_{(e)} \rightarrow (x_{(e),P} \wedge x_{(e),W} \wedge x_{(e),R}) \neg x_{(e),F}))
\end{aligned}
$$

$$
\begin{aligned}
&\wedge (z_W \rightarrow (x_{(a),W} \wedge x_{(b),W} \wedge x_{(c),W} \wedge x_{(d),W} \wedge x_{(e),W}))\\
&\wedge (\neg z_W \rightarrow (\neg x_{(a),W} \vee \neg x_{(b),W} \vee \neg x_{(c),W} \vee \neg x_{(d),W} \vee \neg x_{(d),W}))\\
&\wedge (z_R \rightarrow (x_{(a),R} \wedge x_{(b),R} \wedge x_{(c),R} \wedge x_{(d),R} \wedge x_{(e),R}))\\
&\wedge (\neg z_R \rightarrow (\neg x_{(a),R} \vee \neg x_{(b),R} \vee \neg x_{(c),R} \vee \neg x_{(d),R} \vee \neg x_{(d),R}))\\
&\wedge (z_F \rightarrow (x_{(a),F} \wedge x_{(b),F} \wedge x_{(c),F} \wedge x_{(d),F} \wedge x_{(e),F}))\\
&\wedge (\neg z_F \rightarrow (\neg x_{(a),F} \vee \neg x_{(b),F} \vee \neg x_{(c),F} \vee \neg x_{(d),F} \vee \neg x_{(d),F}))\\
&\wedge (z_P \rightarrow (x_{(a),P} \wedge x_{(b),P} \wedge x_{(c),P} \wedge x_{(d),P} \wedge x_{(e),P}))\\
&\wedge (\neg z_P \rightarrow (\neg x_{(a),P} \vee \neg x_{(b),P} \vee \neg x_{(c),P} \vee \neg x_{(d),P} \vee \neg x_{(d),P}))
\end{aligned}
$$

where the variables corresponds to the focal elements determined by items (a), (b), (c), (d), and (e) and the elements of $S = \{W, R, F, P\}$. From this formula, by using some suitable coding package to translate propositional formulas into a DNNF circuit, the air traffic control support software would be ready to receive input for specific cases.

Let us imagine there is a flight scheduled to land during a period of changing weather conditions, and the air traffic control support system collects the following basic belief assignments:

$$
m_1(A) = \begin{cases} 0.75 & \text{if } A = \{W, R\},\\ 0.25 & \text{if } A = S,\\ 0 & \text{otherwise.} \end{cases}
$$

$$
m_2(B) = \begin{cases} 0.6 & \text{if } B = \{F, P\},\\ 0.4 & \text{if } B = S,\\ 0 & \text{otherwise.} \end{cases}
$$

$$
m_2(C) = \begin{cases} 0.8 & \text{if } C = \{R, F\},\\ 0.2 & \text{if } C = S,\\ 0 & \text{otherwise.} \end{cases}
$$

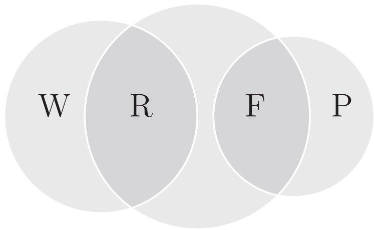


Figure 6.7: Representation of the example's input by a Venn diagram. The size of the area represents the degree of certainty of the corresponding piece of evidence.

The corresponding weight function for this input would be:

$$w(l) = \begin{cases} 0.75 & \text{if } l = y_{(a)}, \\ 0.25 & \text{if } l = \neg y_{(a)}, \\ 0.6 & \text{if } l = y_{(b)}, \\ 0.4 & \text{if } l = \neg y_{(b)}, \\ 0.8 & \text{if } l = y_{(c)}, \\ 0.2 & \text{if } l = \neg y_{(c)}, \\ 1 & \text{otherwise.} \end{cases}$$

Now, by running weighted model counting on the available DNNF circuit according to the previous weights, air traffic controllers can quickly compute the degree of belief for propositions of interest. For example, they may be interested in the degree of belief for 'anomaly due to sub-optimal runway condition', represented by the set $\{R\}$, to suggest diverting the flight. Or in the combined mass function of 'anomaly due to sub-optimal runway condition or failure in the flight control sensors', represented by $\{R, F\}$ to plan an emergency landing. Another proposition of interest might be 'anomaly due to sub-optimal runway condition or sub-optimal pilot performance', represented by $\{R, P\}$ to ask the pilot to turn around and try again. In the current example, these values are $\text{bel}(\{R\}) = 0.43$, $m(\{R, F\}) = 0, 8$ and $m(\{R, P\}) = 0.43$, respectively.

The described example satisfies all the conditions from 1 to 6, which have been introduced as ideal conditions for using this method. For instance, the same DNNF circuit can be used for different flights or for the same flight at different times, since the evidence potentially collected corresponds to the same variables. Additionally, last minute changes in the evidence can be made. For example, if the weather service calls to say that the expected storm will be delayed for an hour, the basic belief assignment m_1 with focal element $\{W, R\}$ must be forgotten. There are various approaches to forget evidence in the Dempster-Shafer context (Kramosil, 1999; Pinto Prieto et al., 2024; Xiaojing et al., 2022), due to the straightforward approach leading to exponential time computations (Smets, 1995). In our case, we can do this efficiently by running a query for a new weight function, in this case

$$w(l) = \begin{cases} 0.6 & \text{if } l = y_{(b)}, \\ 0.4 & \text{if } l = \neg y_{(b)}, \\ 0.8 & \text{if } l = y_{(c)}, \\ 0.2 & \text{if } l = \neg y_{(c)}, \\ 1 & \text{otherwise.} \end{cases}$$

This example is inspired by (Garcia et al., 2024), where the authors apply Dempster-Shafer theory to study real air traffic data. Other areas that could benefit from this approach for similar reasons as the current example are applications of the Internet of Things (Hamda et al., 2023) and intrusion detection systems (Qiu et al., 2022).

Chapter 7

Conclusion

Throughout this dissertation, we have studied the problem of combining evidence that may be partial, mutually contradictory, and with varying degrees of certainty. During this process, we have defined new mathematical, computational and conceptual tools that enhance our understanding of the subject. We summarize the highlights of these contributions here, along with a list of future research directions.

Chapter 2: Background Apart from introducing the required elements of Dempster-Shafer theory and topological models of evidence to follow this dissertation, in this chapter we presented:

1. A theoretical motivation for combining these two frameworks. The key points are that both frameworks use set representation of evidence; aim to compute belief based on evidence; and introduce an intermediate step between the basic evidence and belief, consisting of generating combined evidence.

2. A discussion on how these two frameworks complement each other. We highlight that Dempster-Shafer theory includes quantitative elements, such as the degree of uncertainty associated with the evidence pieces and graded belief. In contrast, topological models of evidence are purely qualitative. However, topological models of evidence define more fine-grained epistemic concepts, such as argument and justification, that may be beneficial for Dempster-Shafer mathematical methods as well.

3. A common vocabulary that allows us to integrate elements from both frameworks into one. In particular, we define S as a set of possible states,

$\boldsymbol{\mathcal{E}}^Q = (E_1, p_1), \dots, (E_\ell, p_\ell)$ as a set of uncertain evidence, where $E_j \subset S$ represents a piece of evidence for every $1 \leq j \leq \ell$ and p_j represents its degree of certainty.

Although this chapter did not present research results itself, the aforementioned comparison may serve as a trailhead for other connections between these two frameworks beyond the ones explored in this work.

Chapter 3: Multi-Layer Belief Model for Combining Uncertain Evidence In this chapter, we developed a new method to compute belief functions (in the Dempster-Shafer sense) that integrates notions from topological models of evidence. Our main contributions on this topic were:

1. Defining a belief model that (*i*) takes inputs also valid for Dempster-Shafer framework, (*ii*) generates a belief function in the Dempster-Shafer sense, and (*iii*) adds parameters to obtain different evidence combinations without the need for switching frameworks or interpretations.

2. Splitting the model into quantitative, qualitative, and bridging components, making their roles explicit. This allows modifications to one layer without compromising the theoretical foundation of the others.

3. Introducing the notion of evidential demand of the agent. This component is represented by the concept of *justification*, which is inspired by the homonymous concept from topological models of evidence. This evidential demand allows distinguishing between the basic evidence accessible to the agent, the arguments supported by this evidence—similarly to topological models of evidence—and the arguments considered convincing enough to generate beliefs of the agent. In our case, we determine justifications by taking into account consistency requirements.

4. Introducing the notion of evidence allocation functions. This element captures the agent's interpretation of having pieces of evidence from different sources simultaneously. The formal definition of these functions is tied to consistency.

5. Introducing the minimum dense set function as an alternative to intersection and union to combine evidence.

6. Showing that, under the restriction of having one piece of evidence per source, the multi-layer belief model can reproduce Dempster's rule of combination, the belief operator of topological models of evidence, and the belief function obtained by the transferable belief model.

7. Proposing two new combination rules. One generated by the multi-layer belief model with Dempster-Shafer frame of justification and union as evi-

dence allocation function. Another one generated by the multi-layer belief model with Dempster-Shafer frame of justification and the minimum dense set function as evidence allocation function. In terms of interpretation, the former can be seen as an alternative to the disjunctive rule of combination (Definition 2.12), and the latter to the unnormalized conjunctive rule of the transferable belief model (Definition 2.11).

These results extend to various research lines:

a) A natural extension would be to define this model for basic belief assignments with arbitrary numbers of focal elements.

b) Additionally, defining a way to add new pieces of evidence or remove old ones without rerunning the whole process from scratch would be valuable. If this is possible, it would be interesting to study the mathematical properties, such as commutativity, associativity, and the existence of a unique neutral element, and compare them with those of other combination rules (Sentz and Ferson, 2002).

c) New frames of justification could be explored. For example, by relaxing the requirement of denseness, such as allowing an element of the topology to intersect the basic evidence but not their intersections (and progressively adding layers of intersections until reaching the strong denseness frame of justification). Another approach could be to require the elements of the topology to intersect *almost all* the elements of the topology, adapting the term "almost all" as necessary. Alternatively, frames of justification could be defined according to a different notion rather than consistency, or on top of consistency, such as requiring large intersections (adapting what "large" means).

d) The δ function can also be modified to capture other families of combination rules. For example, those defined by Denœux (2008) for dependent sources.

e) The established link between the multi-layer belief model and topological models of evidence can also be used to define a quantitative version of topological models of evidence, by introducing graded belief and graded certainty.

Chapter 4: A Qualitative Logic for Evidence and Belief Comparison
In this chapter, we proposed a qualitative logic to compare the strength of belief and evidential support according to evidence certainty. Our main contributions in this regard are:

1. Defining an order among propositions based on the certainty degree of their evidential support. This order only considers certainty values; there are no additional pieces of evidence or set operations such as intersection or union

involved. This is achieved by defining an order-lifting from the natural order among the pieces of evidence according to their certainty value.

2. Justifying the suitability of such certainty order among propositions and connection with Egli-Milner order (Shi et al., 2017) and min-max extension (Maly, 2020) in the finite case.

3. Proposing a qualitative logic that connects evidence and belief based on quantitative models. Notice that the graded belief represented in the models corresponds to the degree of belief obtained by applying a combination method such as Dempster's rule of combination.

4. Collecting validities and invalidities of this logic.

5. A comparison with three naturally related logics, defined in (Ghosh and De Jongh, 2013; Harmanec and Hájek, 1994; Shi et al., 2017). We show that our logic contains the qualitative belief logic of Harmanec and Hájek (1994), partially contains the $KD45-O$ logic of Ghosh and De Jongh (2013), and is strictly different from the logic for convex order of Shi et al. (2017).

Given the exploratory nature of this chapter, the next steps would serve to further consolidate the listed results:

a) Finding a complete axiomatization of the defined logic.

b) Establishing stronger connections, if they exist, between evidential and belief operators.

c) Identifying validities that determine the exact parameter setting (selected frame of justification and evidence allocation function) of the multi-layer belief model. That is, defining more specific models and understanding the differences between the respective logics.

d) Introducing some notion of justification in the logic and evaluating how the logic changes.

Chapter 5: In this chapter, we initiated a systematic computational complexity analysis for Dempster's rule of combination and the multi-layer belief model. Our main contributions are:

1. Finding that restricting the evidence to basic belief assignments with a single proper focal element still makes the problem of applying Dempster's rule of combination #P-complete, even when computing the degree of belief for a singleton. The only restriction that makes the problem polynomial is when basic belief assignments have proper focal elements of size smaller than or equal to c, which is too restrictive in practice.

2. Proposing three ways for combining arbitrary evidence by using the algorithm defined in (Shafer and Logan, 2008) for efficiently applying Dempster's rule of combination on hierarchical evidence. The first method involves assessing whether an arbitrary evidence set is hierarchical, which can be done in polynomial time, and applying Shafer and Logan's algorithm if it is. The second method covers filtering an arbitrary body of evidence to obtain the largest hierarchy contained in it, which is a fixed-parameter tractable problem for a suitable parameter. The third method is an algorithm that merges arbitrary evidence by using Shafer and Logan's algorithm in some key parts and is fixed-parameter tractable.

3. Defining an efficient algorithm to compute the outcome of the belief operator of topological models of evidence.

4. Showing that applying the multi-layer belief model in its most general form—i.e., without specific restrictions on frame of justification or evidence allocation function—is a #P-complete problem. In other words, the expressivity that this method adds for combining evidence, compared to Dempster's rule of combination, does not increase its computational complexity.

These results could be extended by:

a) Further developing the fixed-parameter alternatives to compute Dempster's rule of combination. Our study initiates a promising exploration, yet there is room to identify better parameters.

b) Implementing the proposed algorithm for the belief operator of topological models of evidence and studying its performance in applications. Note that this operator is Boolean and the evidence of the input has all the same certainty.

c) Analyzing the computational complexity of the multi-layer belief model in specific cases beyond Dempster's rule of combination and topological models of evidence. Exploring fixed-parameter tractable alternatives, as with Dempster's rule.

Chapter 6: Knowledge Compilation for Combining Uncertain Evidence This chapter presented a concrete algorithmic solution to bring Dempster's rule of combination to real-life applications without incurring the cost of its intrinsic computational complexity (in suitable contexts). Our key contributions include:

1. Defining a propositional formula that establishes an equivalence between Dempster's rule of combination (and its unnormalized version) for simple basic belief assignments and weighted model counting.

2. Identifying compilation languages known for answering weighted model counting queries in polynomial time, and concluding that Dempster's rule of combination (and its unnormalized version) can be computed efficiently after a suitable compilation to one of these languages.

3. Finding an $\mathrm{OBDD}_<$ representation that makes the whole process, including the compilation part, efficient when the evidence has hierarchical structure.

Although the results of this chapter are somewhat conclusive, they could be extended by applying these techniques in practical situations and comparing their convenience with respect to other alternatives, both in terms of algorithm and theoretical framework.

In summary, this research has elucidated to some extent the problem of combining uncertain evidence, emphasizing both its logical and computational complexity aspects. With this dissertation, we aim to contribute not only to a better understanding of this problem, but also to inspire further research.

Index

List of Symbols

S	set of possible states
2^S	power set of S
P	proposition
m	basic belief assignment
Bel	belief function
$\mathrm{Bel}_{\mathcal{J}}(f,\cdot)$	multi-layer belief function
bel	Dempster-Shafer belief function
plau	Dempster-Shafer plausibility function
com	Dempster-Shafer commonality function
$\oplus$	Dempster's rule of combination
$\boxplus$	unnormalized rule of combination
$\Cup$	disjunctive rule of combination
$\boldsymbol{\mathcal{E}}$	evidence set
E	piece of evidence
$2^{\boldsymbol{\mathcal{E}}}$	power set of $\boldsymbol{\mathcal{E}}$
$\boldsymbol{E}$	element of $2^{\boldsymbol{\mathcal{E}}}$
$\boldsymbol{\tau}_{\boldsymbol{\mathcal{E}}}$	topology generated by $\boldsymbol{\mathcal{E}}$
$\boldsymbol{\mathcal{E}}^Q$	quantitative evidence set
(E,p) or (E,q)	quantitative piece of evidence
$\boldsymbol{\mathcal{J}}$	frame of justification
$\boldsymbol{\mathcal{J}}^{DS}$	Dempster-Shafer frame of justification
$\boldsymbol{\mathcal{J}}^{SD}$	strong-denseness frame of justification
$(Q,\leq)$	order among certainty values
$(\boldsymbol{\mathcal{E}},\leq_e)$	order among evidence pieces
$(2^{\boldsymbol{\mathcal{E}}},\preceq_e)$	order lifting from $(\boldsymbol{\mathcal{E}},\leq_e)$
$(2^S,\trianglelefteq_e)$	order among propositions
$\trianglelefteq_e$	evidence comparison operator
$\trianglelefteq_b$	belief comparison operator
$\mathcal{M}$	evidence-belief model
$\mathcal{H}$	hierarchy of focal elements

$\mathcal{F}$	set of focal elements
$\rightleftharpoons$	conflict between pairs of focal elements
$\overset{r}{\rightleftharpoons}$	r conflicts between pairs of focal elements
$\sim_{\mathcal{E}}$	identical support relation
$S/{\sim_{\mathcal{E}}}$	quotient set of S
$[a]$	quotient class
$WMC(\varphi, w)$	weighted model counting of φ with weights w
$\mathcal{V}$	set of propositional variables
$x \rightarrow^{*} y$	diagram branch from x to y

Bibliography

Arora, S., and Barak, B. (2009). *Computational complexity: A modern approach*. Cambridge University Press.

Aspvall, B., Plass, M. F., and Tarjan, R. E. (1979). A linear-time algorithm for testing the truth of certain quantified Boolean formulas. *Information Processing Letters*, *8*(3), 121–123.

Atitallah, S. B., Driss, M., Boulila, W., Koubaa, A., and Ben Ghézala, H. (2022). Fusion of convolutional neural networks based on Dempster-Shafer theory for automatic pneumonia detection from chest X-ray images. *International Journal of Imaging Systems and Technology*, *32*(2), 658–672. https://doi.org/10.1002/ima.22510

Baltag, A., Bezhanishvili, N., Özgün, A., and Smets, S. (2016). Justified belief and the topology of evidence. In J. A. Väänänen, Å. Hirvonen, and R. J. G. B. de Queiroz (Eds.), *Proceedings of the 23rd international workshop on logic, language, information, and computation (WoLLIC 2016)* (pp. 83–103, Vol. 9803). Springer. https://doi.org/10.1007/978-3-662-52921-8_6

Baltag, A., Bezhanishvili, N., Özgün, A., and Smets, S. (2022). Justified belief, knowledge, and the topology of evidence. *Synthese*, *200*(6), 1–51. https://doi.org/10.1007/s11229-022-03967-6

Baltag, A., Gierasimczuk, N., Özgün, A., Vargas Sandoval, A. L., and Smets, S. (2019a). A dynamic logic for learning theory. *Journal of Logical and Algebraic Methods in Programming*.

Baltag, A., Özgün, A., and Vargas Sandoval, A. L. (2019b). The logic of AGM learning from partial observations. *Proceedings of the 2nd International Workshop on Dynamic Logic (DaLi 2019)*.

Barberà, S., Bossert, W., and Pattanaik, P. K. (2004). Ranking sets of objects. In S. Barberà, P. J. Hammond, and C. Seidl (Eds.), *Handbook of utility theory: Volume 2 extensions* (pp. 893–977). Springer US. https://doi.org/10.1007/978-1-4020-7964-1_4

Barberà, S., and Pattanaik, P. K. (1984). Extending an order on a set to the power set: Some remarks on Kannai and Peleg's approach. *Journal of Economic Theory*, *32*(1), 185–191. https://doi.org/10.1016/0022-0531(84)90083-8

Barnett, J. A. (1981). Computational methods for a mathematical theory of evidence. *Proceedings of the 7th International Joint Conference on Artificial Intelligence (IJCAI 1981)*, 868–875.

Baroni, P., Gabbay, D., Giacomin, M., and van der Torre, L. (Eds.). (2018). *Handbook of formal argumentation*. College Publications.

van Benthem, J. (1987). Verisimulitude and conditionals. In T. A. F. Kuipers (Ed.), *What is closer to the truth? A parade of approaches to truthlikeness* (pp. 103–128). Rodopi.

van Benthem, J., Fernández-Duque, D., and Pacuit, E. (2012). Evidence logic: A new look at neighborhood structures. *Proceedings of the 9th conference on Advances in Modal Logic (AiML 2012)*.

van Benthem, J., Fernández-Duque, D., and Pacuit, E. (2014). Evidence and plausibility in neighborhood structures. *Annals of Pure and Applied Logic*, *165*(1), 106–133. https://doi.org/10.1016/j.apal.2013.07.007

van Benthem, J., Girard, P., and Roy, O. (2009). Everything else being equal: A modal logic for ceteris paribus preferences. *Journal of Philosophical Logic*, *38*(1), 83–125. https://doi.org/10.1007/s10992-008-9085-3

van Benthem, J., and Pacuit, E. (2011). Dynamic logics of evidence-based beliefs. *Studia Logica*, *99*(1-3), 61–92. https://doi.org/10.1007/s11225-011-9347-x

Bergsten, U., and Schubert, J. (1993). Dempster's rule for evidence ordered in a complete directed acyclic graph. *International Journal of Approximate Reasoning*, *9*(1), 37–73. https://doi.org/10.1016/0888-613X(93)90006-Y

Blackburn, P., de Rijke, M., and Venema, Y. (2001). *Modal logic* (Vol. 53). Cambridge University Press.

van der Bles, A. M., van der Linden, S., Freeman, A. L. J., Mitchell, J., Galvao, A. B., Zaval, L., and Spiegelhalter, D. J. (2019). Communicating uncertainty about facts, numbers and science. *Royal Society Open Science*, *6*(181870). https://doi.org/10.1098/rsos.181870

Brink, C. (1992). *Power structures and their application* [Doctoral dissertation, University of Johannesburg].

Chaki, J. (2023). Applications of different methods to handle uncertainty in artificial intelligence. In *Handling uncertainty in artificial intelligence* (pp. 93–101). Springer Nature Singapore. https://doi.org/10.1007/978-981-99-5333-2_7

Chen, J., Kanj, I. A., and Xia, G. (2010). Improved upper bounds for vertex cover. *Theoretical Computer Science*, *411*(40), 3736–3756. https://doi.org/10.1016/j.tcs.2010.06.026

Corsi, E. A., Flaminio, T., Godo, L., and Hosni, H. (2023). A modal logic for uncertainty: A completeness theorem. *13th International Symposium on Imprecise Probabilities: Theories and Applications (ISIPTA 2023)*, (215), 119–129.

Darwiche, A., and Marquis, P. (2002). A knowledge compilation map. *J. Artif. Int. Res.*, *17*(1), 229–264.

Dempster, A. P. (1967). Upper and lower probabilities induced by a multivalued mapping. *The Annals of Mathematical Statistics*, *38*(2), 325–339. https://doi.org/10.1214/aoms/1177698950

Denœux, T. (2008). Conjunctive and disjunctive combination of belief functions induced by nondistinct bodies of evidence. *Artificial Intelligence*, *172*(2), 234–264. https://doi.org/10.1016/j.artint.2007.05.008

Destercke, S., and Burger, T. (2013). Toward an axiomatic definition of conflict between belief functions. *IEEE transactions on cybernetics*, *43*(2), 585–596. https://doi.org/10.1109/TSMCB.2012.2212703

Ding, Y., Holliday, W. H., and Icard, T. F. (2021). Logics of imprecise comparative probability. *International Journal of Approximate Reasoning*, *132*, 154–180. https://doi.org/10.1016/j.ijar.2021.02.004

Dubois, D., Godo, L., and Prade, H. (2023). An elementary belief function logic. *Journal of Applied Non-Classical Logics*, *33*(3-4), 582–605. https://doi.org/10.1080/11663081.2023.2244366

Dubois, D., and Prade, H. (1986). A set-theoretic view of belief functions: Logical operations and approximations by fuzzy sets. *International Journal of General Systems*, *12*(3), 193–226. https://doi.org/10.1080/03081078608934937

Dubois, D., and Prade, H. (1992). On the combination of evidence in various mathematical frameworks. In J. Flamm and T. Luisi (Eds.), *Reliability data collection and analysis* (pp. 213–241). Springer Netherlands. https://doi.org/10.1007/978-94-011-2438-6_13

Dugundji, J. (1965). *Topology*. Prentice Hall.

Ebbinghaus, H.-D., Jörg, F., and Wolfgang, T. (2018). *Mathematical logic*. Springer.

Fine, K. (2017). A theory of truthmaker content I: Conjunction, disjunction and negation. *Journal of Philosophical Logic*, *46*, 625–674.

Flaminio, T., Godo, L., and Ugolini, S. (2022). An approach to inconsistency-tolerant reasoning about probability based on Łukasiewicz logic. In F. Dupin de Saint-Cyr, M. Öztürk-Escoffier, and N. Potyka (Eds.), *Scalable uncertainty management* (pp. 124–138). Springer International Publishing.

Garcia, E. J., Beres, S. L., Mulvihill, M. L., Stephens, C., and Napoli, N. J. (2024). Multiclass flight anomaly detection using Dempster-Shafer theoretic sensor fusion. *AIAA SCITECH 2024 Forum*. https://doi.org/10.2514/6.2024-2615

Ghosh, S., and de Jongh, D. (2013). Comparing strengths of beliefs explicitly. *Logic Journal of the IGPL*, *21*(3), 488–514. https://doi.org/10.1093/jigpal/jzs050

Gierasimczuk, N. (2010). *Knowing one's limits: Logical analysis of inductive inference* [Doctoral dissertation, University of Amsterdam].

Gordon, J., and Shortliffe, E. H. (2008). A method for managing evidential reasoning in a hierarchical hypothesis space. *Yager R.R., Liu L. (eds) Classic Works of the Dempster-Shafer Theory of Belief Functions. Studies in Fuzziness and Soft Computing*, *219*, 868–875. https://doi.org/10.1007/978-3-540-44792-4_12

Greenhill, C. (2000). The complexity of counting colourings and independent sets in sparse graphs and hypergraphs. *Computational Complexity*, *9*(1), 52–72.

Haenni, R. (2005). Shedding new light on Zadeh's criticism of Dempster's rule of combination. *Proceedings of the 7th International Conference on Information Fusion (FUSION 2005)*, *2*, 6. https://doi.org/10.1109/ICIF.2005.1591951

Halpern, J. Y. (2017). *Reasoning about uncertainty*. The MIT Press.

Hamda, N., Hadjali, A., and Lagha, M. (2023). Multisensor data fusion in IoT environments in Dempster-Shafer theory setting: An improved evidence distance-based approach. *Sensors*, *23*, 5141. https://doi.org/10.3390/s23115141

Harmanec, D., and Hájek, P. (1994). A qualitative belief logic. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, *02*(02), 227–236. https://doi.org/10.1142/S0218488594000171

Herawati, L., and Eviyanti, A. (2021). Gynecological disease diagnosis expert system using the web-based Dempster-Shafer method. *PELS*, *1*(2).

Hintikka, K. J. J. (1962). *Knowledge and belief: An introduction to the logic of the two notions*. Cornell University Press.

Hung, N. D. (2017). A generalization of probabilistic argumentation with Dempster-Shafer theory. In G. Kern-Isberner, J. Fürnkranz, and M. Thimm (Eds.), *Proceedings of the 40th annual german conference on advances in artificial intelligence (KI 2017)* (pp. 155–169). Springer International Publishing.

Kannai, Y., and Peleg, B. (1984). A note on the extension of an order on a set to the power set. *Journal of Economic Theory*, *32*(1), 172–175. https://doi.org/10.1016/0022-0531(84)90080-2

Kelly, K. T. (1996). *The logic of reliable inquiry*. Oxford, England: Oxford University Press USA.

Kelly, K. T., and Lin, H. (2021). Beliefs, probabilities, and their coherent correspondence. In I. Douven (Ed.), *Lotteries, knowledge, and rational belief: Essays on the lottery paradox* (pp. 185–222). Cambridge University Press.

Kimmig, A., van Broeck, G., and de Raedt, L. (2017). Algebraic model counting. *Journal of Applied Logic*, *22*, 46–62. https://doi.org/10.1016/j.jal.2016.11.031

Kramosil, I. (1999). Measure-theoretic approach to the inversion problem for belief functions. *Fuzzy Sets and Systems*, *102*(3), 363–369.

Kusuma, R., and Nas, C. (2023). Expert system to diagnose mental health disorders using the Dempster-Shafer algorithm. *Journal of Information Systems and Informatics*, *5*(1), 391–406. https://doi.org/10.51519/journalisi.v5i1.461

Lefevre, E., Colot, O., and Vannoorenberghe, P. (2002). Belief function combination and conflict management. *Information Fusion*, *3*(2), 149–162. https://doi.org/10.1016/S1566-2535(02)00053-2

Maly, J. (2020). *Ranking sets of objects: How to deal with impossibility results* [Doctoral dissertation, Technische Universität Wien]. 10.34726/hss.2020.83187

Murphy, C. K. (2000). Combining belief functions when evidence conflicts. *Decision Support Systems*, *29*(1), 1–9. https://doi.org/10.1016/S0167-9236(99)00084-6

Ogiwara, M., and Hemachandra, L. A. (1993). A complexity theory for feasible closure properties. *Journal of Computer and System Sciences*, *46*(3), 295–325.

Orponen, P. (1990). Dempster's rule of combination is #P-complete. *Artificial Intelligence*, *44*(1), 245–253. https://doi.org/10.1016/0004-3702(90)90103-7

Özgün, A. (2017). *Evidence in epistemic logic: A topological perspective* [Doctoral dissertation, Univerisity of Amsterdam].

Pacuit, E. (2017). *Neighborhood semantics for modal logic*. Springer Publishing Company, Incorporated.

Pearl, J. (1990). Reasoning with belief functions: An analysis of compatibility. *International Journal of Approximate Reasoning*, *4*(5-6), 363–389.

Pichon, F., and Denoeux, T. (2010). The unnormalized Dempster's rule of combination: A new justification from the Least Commitment Principle and some extensions. *Journal of Automated Reasoning*, *45*(1), 61–87. https://doi.org/10.1007/s10817-009-9152-7

Pinto Prieto, D., and de Haan, R. (2022). Using hierarchies to efficiently combine evidence with Dempster's rule of combination. In J. Cussens and K. Zhang (Eds.), *Proceedings of the 38th conference on uncertainty in artificial intelligence (UAI 2022)* (pp. 1634–1643, Vol. 180). PMLR.

Pinto Prieto, D., and de Haan, R. (2024). *Knowledge compilation for combining uncertain evidence* [Unpublished Manuscript].

Pinto Prieto, D., de Haan, R., and Destercke, S. (2024). How to efficiently decombine belief functions? [To appear]. *Proceedings of the 20th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2024)*.

Pinto Prieto, D., de Haan, R., and Özgün, A. (2023). A belief model for conflicting and uncertain evidence: Connecting Dempster-Shafer theory and the topology of evidence. *Proceedings of the 20th International Conference on Principles of Knowledge Representation and Reasoning (KR 2023)*, 552–561. https://doi.org/10.24963/kr.2023/54

Pinto Prieto, D., and Özgün, A. (2024). *A qualitative logic for evidence and belief comparison* [Unpublished Manuscript].

Plotkin, G. D. (1976). A power domain construction. *SIAM Journal on Computing*, *5*(3), 452–487.

Qiu, W., Ma, Y., Chen, X., Yu, H., and Chen, L. (2022). Hybrid intrusion detection system based on Dempster-Shafer evidence theory. *Computers & Security*, *117*, 102709. https://doi.org/10.1016/j.cose.2022.102709

Roman, S. (2008). *Lattices and ordered sets*. Springer New York.

Ruspini, E. H. (1987). The logical foundations of evidential reasoning (revised).

Sentz, K., and Ferson, S. (2002, April). *Combination of evidence in Dempster-Shafer theory* (tech. rep. No. 800792). Sandia National Laboratories. https://doi.org/10.2172/800792

Shafer, G. (1976). *A mathematical theory of evidence*. Princeton University Press.

Shafer, G., and Logan, R. (2008). Implementing Dempster's rule for hierarchical evidence. *Yager R.R., Liu L. (eds) Classic Works of the Dempster-Shafer Theory of Belief Functions. Studies in Fuzziness and Soft Computing*, *219*, 868–875. https://doi.org/10.1007/978-3-540-44792-4_18

Shafer, G., Shenoy, P. P., and Mellouli, K. (1987). Propagating belief functions in qualitative Markov trees. *International Journal of Approximate Reasoning*, *1*(4), 349–400. https://doi.org/10.1016/0888-613X(87)90024-7

Shi, C., Smets, S., and Velázquez-Quesada, F. R. (2017). Argument-based belief in topological structures. *Theoretical Aspects of Rationality and Knowledge*.

Shi, C., and Sun, Y. (2021). Logic of convex order. *Studia Logica*, *109*(5), 1019–1047. https://doi.org/10.1007/s11225-020-09940-z

Smets, P. (1993). Belief functions: The disjunctive rule of combination and the generalized Bayesian theorem. *International Journal of Approximate Reasoning*, *9*(1), 1–35. https://doi.org/10.1016/0888-613X(93)90005-X

Smets, P. (1995). The canonical decomposition of a weighted belief. *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI 2015)*.

Smets, P. (2007). Analyzing the combination of conflicting belief functions. *Information Fusion*, *8*(4), 387–412. https://doi.org/10.1016/j.inffus.2006.04.003

Smets, P., and Kennes, R. (1994). The transferable belief model. *Artificial Intelligence*, *66*(2), 191–234. https://doi.org/10.1016/0004-3702(94)90026-4

Spaans, J., and Doder, D. (2023). Graduality in probabilistic argumentation frameworks. In K. Gal, K. Gal, A. Nowe, G. J. Nalepa, R. Fairstein, and R. Radulescu (Eds.). IOS Press. https://doi.org/10.3233/FAIA230515

Tong, Z., Xu, P., and Denœux, T. (2019). ConvNet and Dempster-Shafer theory for object recognition. *Proceedings of the 13th International Conference on Scalable Uncertainty Management (SUM 2019)*. https://doi.org/10.1007/978-3-030-35514-2_27

Tong, Z., Xu, P., and Denœux, T. (2021). An evidential classifier based on Dempster-Shafer theory and deep learning. *Neurocomputing*, *450*, 275–293. https://doi.org/10.1016/j.neucom.2021.03.066

Troelstra, A. S., and van Dalen, D. (1988). *Constructivism in mathematics: An introduction*. North-Holland.

Valiant, L. G. (1979). The complexity of enumeration and reliability problems. *SIAM Journal on Computing*, *8*(3), 410–421. https://doi.org/10.1137/0208032

Vargas Sandoval, A. L. (2020). *On the path to the truth: Logical & computational aspects of learning* [Doctoral dissertation, University of Amsterdam].

Verbert, K., Babuška, R., and de Schutter, B. (2017). Bayesian and Dempster-Shafer reasoning for knowledge-based fault diagnosis: A comparative study. *Engineering Applications of Artificial Intelligence*, *60*, 136–150. https://doi.org/10.1016/j.engappai.2017.01.011

Vickers, S. (1989). *Topology via logic*. Cambridge University Press.

Wong, S., Yao, Y., Bollmann, P., and Burger, H. (1991). Axiomatization of qualitative belief structure. *IEEE Transactions on Systems, Man, and Cybernetics*, *21*(4), 726–734. https://doi.org/10.1109/21.108290

Xiaojing, F., Deqiang, H., Yi, Y., and Dezert, J. (2022). De-combination of belief function based on optimization. *Chinese Journal of Aeronautics*, *35*(5), 179–193.

Yager, R. R. (1987). On the Dempster-Shafer framework and new combination rules. *Information Sciences*, *41*(2), 93–137. https://doi.org/10.1016/0020-0255(87)90007-7

Yager, R. R., and Liu, L. (2008). *Classic works of the Dempster-Shafer theory of belief functions*. Springer.

Zadeh, L. A. (1986). A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination. *AI Magazine*, *7*(2), 85–90.

Samenvatting

Deze dissertatie gaat in op de uitdaging van het combineren van onzeker, gedeeltelijk en mogelijkerwijs onderling tegenstrijdig bewijs, met een focus op logica en computationele complexiteit. In deze context wordt *bewijs* gerepresenteerd door een deelverzameling van een verzameling met mogelijkheden; *onzekerheid* door waarden tussen 0 en 1; *gedeeltelijkheid* door onwetendheid over deelverzamelingen die niet als bewijs worden gepresenteerd, en *onderlinge tegenstrijdigheid* door een lege doorsnede. Het hoofdprobleem is (1) hoe zulk bewijs gecombineerd kkan worden om een genormaliseerd geheel te krijgen waar de zekerheidswaarden samen opgeteld 1 opleveren. Een relevante uitbreiding is (2) om graden van geloof te berekenen, gebaseerd op dit gecombineerde bewijs. We verkennen problemen (1) en (2) via drie verschillende aanpakken.

Ten eerste bestuderen we formele methoden die (1) en (2) aanpakken. In Hoofdstuk 2 introduceren we wat oplossingen uit Dempster-Shafer-theorie—dat gebaseerd is op het verruimen van axioma's uit de waarschijnlijkheidsleer—en topologische modellen van bewijs—modellen voor epistemische logica die gebaseerd zijn op topologische semantiek. We stellen een gemeenschappelijk vocabulair vast tussen beide kaders, en gebruiken ze als de basis voor onze oplossing voor (1) en (2), dat we uitwerken in Hoofdstuk 3.

Ten tweede presenteren we een modale logica, in Hoofdstuk 4, om proposities te vergelijken in termen van (a) hun graad van geloof (op basis van het model uit Hoofdstuk 3), en (b) de zekerheid van hun bewijskrachtige ondersteuning. De syntaxis van de logica bevat een binaire modale operator voor elke soort vergelijking. De semantiek van de logica maakt gebruik van volgordeverheffingen om (b) te interpreteren.

Ten slotte analyseren we de computationele complexiteit van (1) en (2). In Hoofdstuk 5 richten we ons op de oplossing voor (1) die Dempsters combinatieregel heet, op de geloofsoperator van topologische modellen van bewijs, en op de formele methode die we ontwikkelden in Hoofdstuk 3. In Hoofdstuk 6 leggen we voor om kenniscompilatietechnieken te gebruiken om Dempsters combinatieregel te berekenen.

Abstract

This dissertation addresses the challenge of combining uncertain, partial and possibly mutually contradictory evidence with a focus on logic and computational complexity. In this context, *evidence* is represented as a subset of a universe; *uncertainty* as values between 0 and 1; *partiality* as ignorance about subsets that are not presented as evidence; and *mutual contradiction* as empty intersection. The main problem is (1) to combine such evidence to obtain a normalized body of evidence, where certainty values sum up to 1. A relevant extension is (2) to compute degrees of belief based on that combined evidence. We explore problems (1) and (2) via three different approaches.

First, we study formal methods that address (1) and (2). In Chapter 2, we introduce some solutions from Dempster-Shafer theory—that relaxes probability axioms—and topological models of evidence—models for epistemic logic based on topological semantics. We establish a common vocabulary between both frameworks and use them as the basis of our solution for (1) and (2), developed in Chapter 3.

Second, we present a modal logic in Chapter 4 to compare propositions in terms of (a) their degree of belief based on Chapter 3, and (b) the certainty of their evidential support. Syntactically, it includes a binary modal operator for each type of comparison. Semantically, we make use of order liftings to interpret (b).

Lastly, we analyze the computational complexity of (1) and (2). In Chapter 5, we focus on the solution for (1) named Dempster's rule of combination, the belief operator of topological models of evidence, and the formal method developed in Chapter 3. In Chapter 6, we propose applying knowledge compilation techniques to compute Dempster's rule of combination.

Titles in the ILLC Dissertation Series:

ILLC DS-2020-03: **Jouke Witteveen**
Parameterized Analysis of Complexity

ILLC DS-2020-04: **Joran van Apeldoorn**
A Quantum View on Convex Optimization

ILLC DS-2020-05: **Tom Bannink**
Quantum and stochastic processes

ILLC DS-2020-06: **Dieuwke Hupkes**
Hierarchy and interpretability in neural models of language processing

ILLC DS-2020-07: **Ana Lucia Vargas Sandoval**
On the Path to the Truth: Logical & Computational Aspects of Learning

ILLC DS-2020-08: **Philip Schulz**
Latent Variable Models for Machine Translation and How to Learn Them

ILLC DS-2020-09: **Jasmijn Bastings**
A Tale of Two Sequences: Interpretable and Linguistically-Informed Deep Learning for Natural Language Processing

ILLC DS-2020-10: **Arnold Kochari**
Perceiving and communicating magnitudes: Behavioral and electrophysiological studies

ILLC DS-2020-11: **Marco Del Tredici**
Linguistic Variation in Online Communities: A Computational Perspective

ILLC DS-2020-12: **Bastiaan van der Weij**
Experienced listeners: Modeling the influence of long-term musical exposure on rhythm perception

ILLC DS-2020-13: **Thom van Gessel**
Questions in Context

ILLC DS-2020-14: **Gianluca Grilletti**
Questions & Quantification: A study of first order inquisitive logic

ILLC DS-2020-15: **Tom Schoonen**
Tales of Similarity and Imagination. A modest epistemology of possibility

ILLC DS-2020-16: **Ilaria Canavotto**
Where Responsibility Takes You: Logics of Agency, Counterfactuals and Norms

ILLC DS-2020-17: **Francesca Zaffora Blando**
Patterns and Probabilities: A Study in Algorithmic Randomness and Computable Learning

ILLC DS-2021-01: **Yfke Dulek**
Delegated and Distributed Quantum Computation

ILLC DS-2021-02: **Elbert J. Booij**
The Things Before Us: On What it Is to Be an Object

ILLC DS-2021-03: **Seyyed Hadi Hashemi**
Modeling Users Interacting with Smart Devices

ILLC DS-2021-04: **Sophie Arnoult**
Adjunction in Hierarchical Phrase-Based Translation

ILLC DS-2021-05: **Cian Guilfoyle Chartier**
A Pragmatic Defense of Logical Pluralism

ILLC DS-2021-06: **Zoi Terzopoulou**
Collective Decisions with Incomplete Individual Opinions

ILLC DS-2021-07: **Anthia Solaki**
Logical Models for Bounded Reasoners

ILLC DS-2021-08: **Michael Sejr Schlichtkrull**
Incorporating Structure into Neural Models for Language Processing

ILLC DS-2021-09: **Taichi Uemura**
Abstract and Concrete Type Theories

ILLC DS-2021-10: **Levin Hornischer**
Dynamical Systems via Domains: Toward a Unified Foundation of Symbolic and Non-symbolic Computation

ILLC DS-2021-11: **Sirin Botan**
Strategyproof Social Choice for Restricted Domains

ILLC DS-2021-12: **Michael Cohen**
Dynamic Introspection

ILLC DS-2021-13: **Dazhu Li**
Formal Threads in the Social Fabric: Studies in the Logical Dynamics of Multi-Agent Interaction

ILLC DS-2021-14: **Álvaro Piedrafita**
On Span Programs and Quantum Algorithms

ILLC DS-2022-01: **Anna Bellomo**
Sums, Numbers and Infinity: Collections in Bolzano's Mathematics and Philosophy

ILLC DS-2022-02: **Jan Czajkowski**
Post-Quantum Security of Hash Functions

ILLC DS-2022-03: **Sonia Ramotowska**
Quantifying quantifier representations: Experimental studies, computational modeling, and individual differences

ILLC DS-2022-04: **Ruben Brokkelkamp**
How Close Does It Get?: From Near-Optimal Network Algorithms to Suboptimal Equilibrium Outcomes

ILLC DS-2022-05: **Lwenn Bussière-Carae**
No means No! Speech Acts in Conflict

ILLC DS-2022-06: **Emma Mojet**
Observing Disciplines: Data Practices In and Between Disciplines in the 19th and Early 20th Centuries

ILLC DS-2022-07: **Freek Gerrit Witteveen**
Quantum information theory and many-body physics

ILLC DS-2023-01: **Subhasree Patro**
Quantum Fine-Grained Complexity

ILLC DS-2023-02: **Arjan Cornelissen**
Quantum multivariate estimation and span program algorithms

ILLC DS-2023-03: **Robert Paßmann**
Logical Structure of Constructive Set Theories

ILLC DS-2023-04: **Samira Abnar**
Inductive Biases for Learning Natural Language

ILLC DS-2023-05: **Dean McHugh**
Causation and Modality: Models and Meanings

ILLC DS-2023-06: **Jialiang Yan**
Monotonicity in Intensional Contexts: Weakening and: Pragmatic Effects under Modals and Attitudes

ILLC DS-2023-07: **Yiyan Wang**
Collective Agency: From Philosophical and Logical Perspectives

ILLC DS-2023-08: **Lei Li**
Games, Boards and Play: A Logical Perspective

ILLC DS-2023-09: **Simon Rey**
Variations on Participatory Budgeting

ILLC DS-2023-10: **Mario Giulianelli**
Neural Models of Language Use: Studies of Language Comprehension and Production in Context

ILLC DS-2023-11: **Guillermo Menéndez Turata**
Cyclic Proof Systems for Modal Fixpoint Logics

ILLC DS-2023-12: **Ned J.H. Wontner**
Views From a Peak: Generalisations and Descriptive Set Theory

ILLC DS-2024-01: **Jan Rooduijn**
Fragments and Frame Classes: Towards a Uniform Proof Theory for Modal Fixed Point Logics

ILLC DS-2024-02: **Bas Cornelissen**
Measuring musics: Notes on modes, motifs, and melodies

ILLC DS-2024-03: **Nicola De Cao**
Entity Centric Neural Models for Natural Language Processing

ILLC DS-2024-04: **Ece Takmaz**
Visual and Linguistic Processes in Deep Neural Networks: A Cognitive Perspective

ILLC DS-2024-05: **Fatemeh Seifan**
Coalgebraic fixpoint logic Expressivity and completeness result

ILLC DS-2024-06: **Jana Sotáková**
Isogenies and Cryptography

ILLC DS-2024-07: **Marco Degano**
Indefinites and their values

ILLC DS-2024-08: **Philip Verduyn Lunel**
Quantum Position Verification: Loss-tolerant Protocols and Fundamental Limits

ILLC DS-2024-09: **Rene Allerstorfer**
Position-based Quantum Cryptography: From Theory towards Practice

ILLC DS-2024-10: **Willem Feijen**
Fast, Right, or Best? Algorithms for Practical Optimization Problems