PhD DISSERTATIONS OF
NORTHWESTERN ITALIAN PHILOSOPHY CONSORTIUM – FINO

Counterfactuals 2.0

Logic, Truth Conditions, and Probability

Giuliano Rosella

2023

Examination Committee/Commissione di esame:

Prof. Jan Sprenger, Prof. Hannes Leitgeb, Prof. Tommaso Flaminio

UNIVERSITÀ DEGLI STUDI DI GENOVA

UNIVERSITÀ DEGLI STUDI DI TORINO

UPO UNIVERSITÀ DEL PIEMONTE ORIENTALE

UNIVERSITÀ DEGLI STUDI DI PAVIA

Counterfactuals 2.0:
Logic, Truth Conditions, and Probability

Giuliano Rosella
Università degli Studi di Torino - (Consorzio FINO)

Supervisor:
Prof. Jan Sprenger

Reviewers:
Prof. Tommaso Flaminio
Prof. Hannes Leitgeb

Contents

Chapter 0

Introduction

0.1 Background and Motivation

On the 20th of April 2023, an article entitled *"The complex math of counterfactuals could help Spotify pick your next favorite song"* was shared by Bill Gates on his Facebook page. The article, written by Will Douglas Heaven for the MIT - Technology Review, explores how statistical frameworks are used to model counterfactual reasoning and how these frameworks can influence personal music suggestions on Spotify through machine learning algorithms.

It is intriguing to observe the growing recognition of counterfactuals in the field of AI, especially considering their origins in classical philosophical discussions. And influential public figures like Bill Gates, who share articles on platforms like Facebook, contribute to the increased attention given to counterfactuals. In the era of exponential technological and data growth, as well as the popularity of Artificial Intelligence (AI), counterfactuals have emerged as a hot topic. They are particularly valuable within the realm of explainable AI, as they provide insights into why AI systems make specific decisions or predictions, which are often difficult to interpret (e.g. see Byrne 2019; Chou, Moreira, et al. 2022; Miller 2019).

Nowadays, the practical applications of counterfactuals have become increasingly prevalent and indispensable. However, it is equally crucial to delve into the theoretical foundations of counterfactuals. By exploring their logic, philosophical underpinnings, and probabilistic interpretation, we can enhance our understanding of how counterfactuals operate, the conditions under which they hold, and their implications for statistical inferences. This investigation would enable the development of more reliable and accurate theoretical frameworks for counterfactual reasoning, thereby improving their practical applications. Furthermore, a comprehensive exploration of counterfactuals on a theoretical level promotes interdisciplinary collaboration among researchers from various disciplines, such as philosophy, computer science, statistics, linguistics, and cognitive science. This collaboration enriches our understanding of counterfactual reasoning from multiple perspectives, leading to more comprehensive and rigorous methodologies.

But before proceeding further, it is essential to clarify what counterfactuals are. In line with the usual practice in philosophy-related fields, counterfactuals are often understood as conditional sentences of the form "If it were the case that [`antecedent`], then [`consequent`] would be the case". For example, "If I hadn't received the vaccination, then I would have caught the flu". Therefore, for our purposes, counterfactuals can be equated with subjunctive conditionals. The basic informal idea behind the evaluation of a counterfactual, as the above-mentioned article[1] puts it,

> [...] is to ask what would have happened in a situation had certain things been different. It's like rewinding the worlds, changing a few crucial details, and then hitting play to see what happens.

Let us consider the example "If I hadn't received the vaccination, then I would have caught the flu"; according to the aforementioned idea, we can conclude that this counterfactual is true. Imagine we rewind the world to the moment before I received the injection and imagine an event preventing the vaccine from entering my body (e.g. a reaction with the substance or accidental spillage). In such a scenario, I would have been unprotected against the virus and would have caught the flu in the following days. Hence, the counterfactual will be evaluated as true. This idea has been formally implemented by Stalnaker (1968) and further developed by Lewis (1973b). The Stalnaker-Lewis account is based upon the idea that a counterfactual "If A were the case, then B would be the case," represented as $A \mathrel{\Box\!\!\rightarrow} B$, is true in a possible world w if and only if B, the consequent, is true in the most similar worlds to w where A, the antecedent, is true. This account is commonly referred to as the "standard account" for counterfactuals. The historical evolution leading to this standard account has passed through the *strict conditional* analysis of counterfactuals and conditional statements in general. It is a consensus in the literature on conditionals that their truth conditions cannot be specified adequately by means of a truth-functional connective. Taking the example from Stalnaker (1968), one may know the truth values of the two sentences "Willie Mays played in the American League" and "Willie Mays hit four hundred", without knowing whether Mays would have hit four hundred if he had played in the American League. A fortiori, conditional sentences, as people ordinarily use them, cannot be formalized by means of the material

[1] Heaven, Will Douglas, "MIT Technology Review", *The complex math of counterfactuals could help Spotify pick your next favorite song*, 4 April 2023.

implication $A \supset B$. Using the material implication as an analysis of conditionals leads to some paradoxical predictions. For instance, whenever a sentence B is true, any conditional having B as its consequence would be predicted as true: $B \models A \supset B$. However, it should not be acceptable to infer that "if John is red-haired, then he is a criminal" from the fact that "John is a criminal" is true.

An alternative analysis of conditional sentences that avoids this kind of paradoxes is due to C. I. Lewis (1912) who put forward a treatment of conditionals in terms of a *strict conditional* operator, represented as $\Box(A \supset B)$, that consists in a necessitated version other the material implication. However, the strict conditional analysis also presents some shortcomings especially when applied to counterfactual conditionals. In particular, the strict conditional validates antecedent strengthening, that is $\Box(A \supset B) \models \Box((A \land C) \supset B)$ and transitivity, $\Box(A \supset B), \Box(B \supset C) \models \Box(A \supset C)$. These two principles do not apply to counterfactual conditionals (see Lewis 1973b): it is acceptable to infer that *if kangaroos had no tails, they would topple over*, but not the antecedent-strengthened version *if kangaroos had no tails, but had four legs, they would topple over*. Analogously, borrowing an example from Stalnaker (1968), it seems acceptable to assert *if J. Edgar Hoover had been born a Russian, then he would have been a Communist* and *If he had been a Communist, he would have been a traitor*; but it is not true that *if he had been born in Russia, he would have been a traitor*. Drawing from similar considerations, Stalnaker (1968) tackled the challenge of formulating an appropriate semantic theory of conditionals that would overcome the problems associated with truth-functional and strict conditional analyses. Stalnaker (1968) proposed that the truth-conditions of a conditional sentence $A > B$ should be explained through a *selection function*. Specifically, $A > B$ is deemed true at a world w whenever B is true in the *unique* world where A is true and which differs minimally from the actual world, as far as A permits it to. The idea of such a world where the antecedent is true and which differs minimally from the world of evaluation is formally implemented by means of a selection f such that, for any proposition X and world w, delivers *the* most similar world to w where X holds, denoted as $f(X, w)$. Stalnaker (1968) acknowledged that this idea was inspired by a suggestion put forth by Ramsey (1931); according to Ramsey's idea, our belief in a conditional sentence such as "if the Chinese enter the Vietnam conflict, the United States will use nuclear weapons" should align with the hypothetical belief in the consequence "the United States will use nuclear weapons" after (hypothetically) adding the antecedent "the Chinese enter the Vietnam conflict" to our stock of beliefs. If the antecedent is inconsistent with our stock of beliefs, Stalnaker argued that we should (hypothetically) make the necessary adjustments to restore consistency. This informal concept mirrors the notion of the most similar world where the antecedent is true. The *most-similar world* analysis has been formally developed by Stalnaker and Thomason (2008) and it overcomes many of the limitations associated with the truth-functional and the strict conditional accounts.

However, it was not until the work of David Lewis (1973) that this style of analysis has been extensively developed. While Stalnaker was committed to a single semantic account of conditionals, Lewis systematically analyzed many different semantic accounts of conditionals, and their associated logic, arising from the various specifications of the *being most similar to* relation. The underlying idea remained consistent: according to Lewis, a conditional $A > B$ is deemed true at w if and only if B is true at the most similar worlds to w where A is true. More precisely, Lewis argued that the relation of *being most similar to*, which is central to the truth-conditions of conditionals, could be specified in various ways, leading to different semantic accounts of conditional sentences. From a more technical perspective, Lewis (1971, 1973b) thoroughly examined the different constraints that one could impose on the relation of *being most similar to* and the logical principles that follow from those constraints. For instance, if we impose that the set of most similar worlds must include at most *one* element, we end up with Stalnaker's account of conditionals. This constraint entails the validity of the principle of *conditional excluded middle*, $\models (A > B) \vee (A > \neg B)$. Later in the thesis, we will delve into Lewis' analysis in more details, discussing the various constraints on the *being most similar to* relation and the characteristic logical principles associated to them. Lewis also coined the term *variably strict conditionals* to refer to the class of conditional sentences he analyzed. The expression was coined to emphasize the contrast with the strict conditional analysis. Specifically, according to Lewis, a variably strict conditional (and in particular a counterfactual) shares some similarities with the strict conditional, as it involves a necessitation of a conditional dependence: a variably strict conditional $A > B$ is true at w if and only if $A \supset B$ is true at *all* the most similar worlds to w that make the antecedent true. However, it differs from the strict conditional in that the set of most similar worlds is variable and dependent on the antecedent. This means that the set of *most similar worlds* is always relative to a given formula, which is the antecedent of the conditional under evaluation. Within the realm of variably strict conditional logics, Lewis (1973b) identified a specific one as his preferred logic of counterfactual conditionals, i.e. logic **VC**. Unlike Stalnaker's account, **VC** invalidates conditional excluded middle; from a semantic perspective, while Stalnaker's account was committed to the idea that the *most similar world* is unique, Lewis counterfactual logic **VC** allowed for the possibility of multiple *most similar* worlds relative a given one. Nevertheless, both

Stalnaker's account and Lewis' counterfactual logic surpass the limitations of the truth-functional and strict conditional analyses.

Furthermore, Lewis also explored various semantic implementations of the *most-similar-world* idea. For instance, Lewis (1971, 1973b) showed how this idea can be implemented using similarity orders or systems of spheres, where worlds are ordered based on their similarity to a given one, or, equivalently, through selection functions, *à la* Stalnaker, taking a formula A and world w as arguments and yielding the set of most similar worlds to w where A is true. Throughout this thesis, we will primarily work within the context of Lewis semantic account(s) of variably strict conditionals. For this reason and, for simplifying the terminology, we may refer to the standard account for counterfactuals as the "Lewis account" instead of Stalnaker-Lewis. However, it is important to acknowledge that Stalnaker's original idea represents the first essential step towards the development of what is now recognized as the standard account.

The standard Lewisian account has faced many criticisms. One critical point is the challenge in defining or characterizing the notion of *similarity*, which is crucial in Lewis' truth conditions for counterfactuals. Some authors in philosophy (e.g. Fine 1975) argue that the use of the similarity relation, due to its vagueness, assigns incorrect truth values to certain counterfactuals; additionally, some scholars from computer science (e.g. Ginsberg 1986), observe that putting into action the notion of similarity in computer programs is challenging, making counterfactuals difficult to implement within AI software.

More recently, the standard Lewisian account has been criticized for its unsatisfactory logical properties. Fine (2012) argues that Lewis' logic of counterfactuals is not the correct one to capture counterfactual reasoning, as it invalidates certain intuitive inferential patterns involving counterfactuals. To address these shortcomings, alternative logical and semantic accounts of counterfactuals based on truthmaker semantics have been proposed (Briggs 2012; Fine 2012).

Another critical point is that the standard account evaluates counterfactuals with impossible antecedents as vacuously true. However, some scholars (e.g. Berto, French, et al. 2017; McLoone 2020; Nolan 1997; Sendłak 2021) believe that certain counterfactuals with impossible antecedents may be false, and they have developed alternative non-classical accounts to address this limitation.

The standard account has been challenged also on a more empirical ground: Ciardelli, Zhang, and Champollion (2018) show that people's performance in counterfactual reasoning does not align with the predictions of the standard account, leading to the proposal of different semantic frameworks based on inquisitive semantics.

Furthermore, interpreting and assigning a probability to counterfactual statements is not clear in the standard Lewisian account. Other frameworks, such as Pearl's (2000) causal modeling semantics, provide better interpretations and probability assignments to counterfactual statements. For instance, in Pearl's framework, the probability of a counterfactual "If it were the case that A, then B would be the case" is characterized as the probability that B holds, under the counterfactual assumption that the antecedent A is the case, i.e. under an intervention on the model that forces A to be true.

These criticisms and alternative proposals highlight how the standard Lewisian theory of counterfactuals, while intuitive and useful for many purposes, does not address all the questions surrounding counterfactual reasoning. Hence, the debate over counterfactuals remains active and flourishing.

The present thesis focuses on the theoretical and logical aspects of Lewis counterfactuals. Although we agree with many of the criticisms just mentioned, rather than rejecting the standard account and proposing an alternative, we aim to *squeeze* and explore the full potential of the standard Lewisian account and address open questions within its framework. The basic motivation behind the present work is to draw the attention to some open questions concerning Lewis counterfactuals that have been neglected, and eventually provide an answer to them. For instance, questions like: *what is the probability of a Lewis counterfactual? Is Lewis counterfactual a definable operator*? *Is it possible to analyze Lewis' logic of counterfactuals from an algebraic perspective*? do not have yet a final answer and they have only be partially addressed in the relevant literature. Specifically, the present thesis intends to shed lights on neglected open problems concerning the logical properties, truth conditions, and probabilistic behavior of Lewis counterfactuals and variably strict-conditionals in general. The analysis will span three levels: the logical behavior within an algebraic setting, philosophical themes on truth conditions and semantics, and the characterization of the probability of Lewis counterfactuals. The investigation will primarily employ a logico-semantic perspective, utilizing methods borrowed from algebraic logic, such as Boolean algebras with operators (Jipsen 1992), Boolean algebras of conditionals (Flaminio, Godo, and Hosni 2020), Jónsson-Tarski duality (Jónsson and Tarski 1951), from philosophical logic, such possible worlds semantics for modal logic (Blackburn, de Rijke, and Venema 2001), sphere-based models for conditionals (Lewis 1971), from formal epistemology, like causal models (Pearl 2000), and from theory of uncertain quantification, such as classical probability functions and Dempster-Shafer belief functions (Dempster 1968). Some among these methods are well-established (e.g. possible worlds semantics, Boolean algebras of operators), some others are rather novel (e.g. Boolean algebras of conditionals) or haven't been applied to classical philosophical themes like Lewis counterfactuals (e.g. Dempster-Shafer theory). Hence, the present work can also be regarded as a test field to show how applications of these novel methods to Lewis counterfactuals contributes to both the understanding of counterfactuals and the advancement of these analytical tools.

0.2 Objectives and Open Problems

In the present thesis, as we mentioned, we aim to address and resolve certain open problems in logic and philosophy related to Lewis counterfactuals and variably strict conditionals in general. These problems are interconnected and fall into three distinct lines of investigation.

0.2.1 Counterfactuals in Algebraic Logic

At the more technical level, we will focus on the following research questions:

Questions - Algebraic Semantics

(A1) Can Lewis' variably strict conditional logics be represented algebraically? This question encompasses several sub-questions:

- (A1a) Which class of algebras, if any, constitutes an equivalent algebraic semantics for Lewis' variably strict conditional logics?
- (A1b) What are the structural properties of these types algebras? For example, how can we characterize the deductive filters within this class?

These questions require further clarification. From a logical standpoint, the ongoing debate on counterfactuals (including variably strict conditionals) lacks a systematic and well-developed algebraic framework to account for their logic. While some preliminary progress has been made by Nute (1975) and Weiss (2019), who introduced certain algebraic structures that serve as *weakly* sound and complete semantics for conditional logics, a detailed algebraic treatment of variably strict conditional logics is still absent. This treatment should, at least, encompass a comprehensive investigation of the corresponding algebras and provide an answer to the question regarding the *algebraizability* of variably strict conditional logics. The aim of our work is to bridge this gap by proposing such an account.

We believe that this algebraic approach would provide valuable insights into the logic of variably strict conditionals. The algebraic methodology is widely recognized as a fruitful approach to delve deeper into the meta-logical properties of a logic system and its connections with other systems; it enables the provision of a semantic interpretation through algebraic structures and sheds new lights on various logico-philosophical problems. For instance, Heyting (1930) and Esakia (2019) offered a semantic interpretation of intuitionistic logic (**IL**) using Heyting algebras, which facilitated a better understanding and proof of certain meta-logical properties of **IL**, such as the finite model property and the absence of a finite truth-table interpretation for **IL**. Moreover, Heyting Algebras played a crucial role in demonstrating the embedding of **IL** into the modal logic **S4** (Gödel 1986; McKinsey and Tarski 1948) and providing a justification for Kripke semantics for **IL**. Algebraic methods have also been employed to investigate some properties of modal logics, including the relationship between the global and the local consequence associated with each normal modal logic (Kracht 1999; Moraschini 2023), and the canonical construction to prove completeness theorems (Blackburn, de Rijke, and Venema 2001). Additionally, algebraic methods have proved useful to provide a semantic interpretation for relevance logics (Font and Rodríguez 1990) and inquisitive logics (Punčochář 2021; Quadrellaro 2022), as well as contributing to the development of a semantic account of paraconsistent belief revision (Carrara, Fazio, and Baldi 2022).

We believe that a systematic algebraic account for variably strict conditional logics will prove equally fruitful in enhancing our understanding of their properties and their relationship with other logical systems. It may also lead, as we will see, to the discovery of new properties or new interpretations of variably strict conditional logics.

0.2.2 Logics and Truth-Conditions of Counterfactuals

From a more philosophical standpoint, we will focus on some less apparent open problems within the standard account. For instance, the following questions arise in the context of the present thesis:

Questions - Counterfactuals in Philosophical Logic

(L1) Can we define Lewis counterfactuals in terms of other conditional operators to achieve a reductionist account of Lewis counterfactuals?

(L1a) What is the relationship between a Stalnaker conditional and a Lewis counterfactual?

(L1b) What is the connection between the probabilistic conditional (a.k.a Adams indicative conditional) and Lewis counterfactuals?

(L1c) Is it possible to have a unified account of Stalnaker conditionals and Lewis counterfactuals? Is it possible to have a unified account of Adams indicative/probabilistic conditionals and Lewis counterfactuals?

(L2) How can we interpret the truth-conditions of a Lewis counterfactual? Can they be interpreted differently from the classical Lewisian interpretation based on similarity among worlds?

Question (L1) has surprisingly been neglected in the literature. To the best of our knowledge, there is no logical framework where the Lewis counterfactual $\Box\!\!\rightarrow$ is a non-primitive operator, nor a definite answer to the question whether Lewis counterfactuals can be defined in terms of Adams or Stalnaker conditionals. However, the idea of defining a counterfactual (or a subjunctive conditional in general) in terms of other conditional operators is not entirely new in literature, although it has not been formally developed, apart from a few exceptions (van Fraassen 1974).

Indeed, a similar idea was suggested by Adams (1975, ch. IV), and then revived by Edgington (2008). They argued that the probability of a subjunctive conditional (and in particular a counterfactual) corresponds to a certain conditional probability, so that a subjunctive conditional turns out to be an indicative conditional in disguise. Here and in the following, we identify indicative conditionals with Adams conditionals, whose probability coincide with the corresponding conditional probability. Adams (1975) considered two hypotheses to explain the inferential behavior of counterfactuals in terms of the corresponding indicative. One hypothesis, referred to as the *prior epistemic probability thesis* by Schulz (2017), is based upon the idea that the probability of a counterfactual corresponds to the probability of a conditional uttered in the past. Thus, what justifies asserting a counterfactual in the present is what justified the past utterance of the corresponding indicative conditional. For example, the utterance "if Oswald hadn't kill Kennedy, someone else would have" would be justified by our past credence in the indicative "if Oswald doesn't kill Kennedy, someone else will". However, Schulz (2017) and Adams (1975) discuss how this view yields unfavorable results, as it predicts the (non-)assertibility and (non-)justification of certain counterfactual judgments that clearly do not hold intuitively[2]. This led Adams to consider an alternative view, which can be seen as a generalization of the prior epistemic probability view, and Schulz (2017) refers to as the *hypothetical epistemic probability view*. Edgington (2008) also embraces a similar view. The idea is that the evaluation and assertion of a counterfactual statement, such as "if it were A, then B would hold" coordinate with the epistemic attitude toward the corresponding indicative conditional "if A is the case, then B will be the case" as uttered in an hypothetical (possibly past) situation. Adams (1975, pp. 129-133) finds problems with this view as well, while Schulz (2017) and Edgington (2008) seem open to the idea that a predictive and satisfactory account of counterfactuals in terms of hypothetical prior probabilities is theoretically possible, albeit challenging to develop. According to this view, the acceptability of the counterfactual "If Oswald hadn't kill Kennedy, someone else would have" is undermined by the relevant hypothetical belief state of evaluation before Kennedy assassination. In that state, we possessed information that only Oswald was prepared to kill the president, ruling out any other option. In this situation, the conditional probability of "someone else will kill Kennedy" given that Oswald doesn't, would be very low (i.e. the indicative "If Oswald doesn't kill Kennedy, someone else will" would not be acceptable according to Adams' account), hereby justifying the unacceptability of the counterfactual "If Oswald hadn't kill Kennedy, someone else would have". Edgington (2008) also provides other examples that can be accounted for through the hypothetical prior probability view. Thus, both Schulz (2017) and Edgington (2008) endorse the plausibility of the hypothetical prior probability, although they acknowledge its difficult implementation, particularly concerning the selection of the relevant hypothetical belief state/situation of evaluation. Overall, it is crucial to emphasize that both the prior epistemic probability view and the hypothetical prior probability view imply, albeit subtly, that counterfactuals are contingent upon or can be defined in terms of their corresponding Adams indicative conditionals.

On a different basis, van Fraassen (1974) argues that a Lewis counterfactual can be defined in terms of a Stalnaker conditional preceded by supervaluationist operator. He demonstrates how Lewis' models for counterfactuals

[2] The Kennedy's assassination example is indeed one of these problematic cases. Suppose before Kennedy's assassination we where aware of a conspiracy intended to kill Kennedy. Under this assumption, we might have accepted the indicative "If Oswald doesn't kill Kennedy, someone else will". Suppose that afterwards we learn new evidence disproving our previous conspiracy theory: we won't be willing to assert "If Oswald hadn't kill Kennedy, someone else would have". This scenario highlights the mismatch between our presumably high past credence in the indicative conditional and the low present credence in the counterfactual, hence providing a counterexample to the prior epistemic probability thesis.

can be derived from a family of Stalnakerian models, presenting a translation theorem from the language of Lewis counterfactuals into the language of Stalnakerian conditional logic equipped with a supervaluationist operator. However, besides this translation result, van Fraassen does not develop a complete logical account based on this new expanded language, for instance he doesn't axiomatize the resulting logic, nor extend the account to the whole family of variably strict conditionals. Except for these attempts (Adams 1965; Edgington 2008; van Fraassen 1974), which are either informal or not fully developed, no conclusive answer has been provided to question (L1).

Therefore, our investigation into question (L1) primarily aims to explore whether Lewis counterfactuals and their logical behavior can be *defined* in terms of Stalnaker conditionals or probabilistic conditionals, a.k.a. indicative conditionals in the sense of Adams. The present work strives to find a unified logical account where both Lewis counterfactuals and Stalnaker conditionals, or Lewis counterfactuals and Adams conditionals, can interact on the same object language level. However, the scope of question (L1) extends beyond the technical level, as we believe it may have significant philosophical implications by offering new enlightening answers to questions (L1a)-(L1b)-(L1c).

Questions (L1a)-(L1b) have been extensively examined in the literature (Adams 1975; Bennett 2003; Edgington 2008; Gibbard 1980; Lewis 1973b; Stalnaker 1968, 1975, 1980) and the connections among Stalnaker conditionals, Adams conditionals and Lewis counterfactuals are well known. For instance, Adams' conditional logic coincides with Stalnaker conditional logic over their common domain (i.e. a language where nested occurrences of the conditional operator are not allowed, see Gibbard 1980); Stalnaker's and Lewis' logic differ in that the former validates the principle of conditional excluded middle, i.e. $\models (A \mathrel{\Box\!\!\rightarrow} B) \vee (A \mathrel{\Box\!\!\rightarrow} \neg B)$, whereas Lewis's counterfactual logic does not (Lewis 1973b). However, these comparisons among Stalnaker, Lewis, and Adams conditionals occur at a meta-level, where we can compare their logics and examine which principles hold. In other words, given a conditional operator in the object language, we can compare how these different theories interpret it and the truth conditions they assign to it. Nevertheless, to the best of our knowledge, no account has been developed where Lewis counterfactuals interact with Stalnaker or Adams conditionals on the same object language level. Consequently, an answer to question (L1c) is still missing. Our work aims to fill this gap by addressing the connections of Lewis counterfactuals with Stalnaker and Adams conditionals, i.e questions (L1a) and (L1b), within the scope of question (L1), and to determine if we can find a unified account for these types of conditionals, providing an answer to question (L1c). The purpose of finding a unified account is to enhance the explanatory power of our philosophical theorizing. It would allow us to explain the differences and connections between Stalnaker conditionals, Adams conditionals, and Lewis counterfactuals within the same language and logical framework. To illustrate this, we present an illuminating example due to van Fraassen (1974). Consider the question:

(1) Would he get life, if he were caught?

Van Fraassen argues that when we ask a question like (1), we typically presuppose conditional excluded middle, i.e. $\models (A \,\Box\!\!\rightarrow B) \vee (A \,\Box\!\!\rightarrow \neg B)$. Consequently, (1) is perceived as a yes-or-on question with two possible answers:

(2) (a) Yes: he would get life if he were caught

(b) No: he would not get life if he were caught

In other words, either "if he were caught he would get life" ($A > B$) and "if he were caught he would not get life" ($A > \neg B$) must hold. By contrast, consider the following two questions:

(3) Is it certain (necessarily, really true) that he would get life if he were caught?

(4) Would he get a life if he were caught, or would he not get life if he were caught?

Van Fraassen holds that it seems clear that (1) fulfills the role of (4), rather than (3). He suggests that

> "...this fact seems inexplicable if we regard the subjunctive conditional [if he were caught, he would get life] as a Lewis conditional (whose negation is 'No, he might not') rather than a Stalnaker conditional (whose negation is 'No, he would not')"(van Fraassen 1974, pp. 188-189)

That is, when we ask question (1), or (4), we presuppose conditional excluded middle, treating the subjunctive conditional "if he were caught, he would get life" as a Stalnaker conditional. However, when we ask question (3), we do not presuppose conditional excluded middle, treating the same subjunctive conditional as a Lewis counterfactuals. Indeed, negating the Lewis counterfactual "if he were caught, he would get life" amounts to asserting that "if he were caught, he might not get a life"[3].

Merely appealing to Stalnaker's theory or Lewis' account is insufficient to distinguish these two different uses of the conditional "if he were caught,

[3]We will see later in the thesis that this feature may be justified by the fact that in Lewis' logic **VC**, the *might*-counterfactual $\Diamond\!\!\rightarrow$ is the dual of the classical *would*-counterfactual, i.e. $A \Diamond\!\!\rightarrow B := \neg(A \,\Box\!\!\rightarrow \neg B)$

he would get life" in the given example. Stalnaker's theory, which requires a conditional to obey conditional excluded middle, cannot account for the failure of conditional excluded middle in the context of question (3). On the other hand, Lewis' account of counterfactuals, in which conditional excluded middle fails, cannot explain why we presuppose conditional excluded middle under the scope of question (1)/(4). This example highlights the need for a unified logical account that incorporates both Stalnaker conditionals and Lewis counterfactuals, as it would enable us to differentiate, on the same language level, between the Stalnakerian and the Lewisian uses of a conditional.

Similarly, it would be desirable to have a uniform account in which Adams probabilistic conditionals, a.k.a. indicative conditionals, interact with Lewis counterfactuals. Consider the classical examples (Adams 1970; Gibbard 1980; Lewis 1973b):

(5) If Oswald didn't kill Kennedy, someone else did

(6) If Oswald hadn't killed Kennedy, someone else would have

According to our ordinary linguistic practice, (5) is considered correctly assertible, or true, while (6) sounds infelicitous. Again, Adams' theory can correctly explain why (5) is assertible, whereas is not sufficient to explain why (6) sounds infelicitous. Likewise, Lewis' theory alone fails to capture this difference, as it correctly predicts the falsity of (6), but fails to account for the truth of (5). The issue arises from the limited expressiveness of Adams' theory and Lewis' theory to accommodate both types of conditionals in (5) and (6). Consequently, a unified logical account of Adams conditionals and Lewis counterfactuals would be more explanatory, allowing us to differentiate between the Adams-like and the Lewisian uses of a conditional at the same language level.

To summarize, while questions (L1a) and (L1b) have been extensively addresses in the existing literature, no unified account of Lewis counterfactuals with Stalnaker conditionals or Adams conditionals have been provided, so addressing question (L1c). In this thesis, we aim to fill this gap by addressing question (L1), which has been neglected in previous research, and striving to develop a unified account of (certain types of) conditionals in which Lewis counterfactuals can be defined.

On the other hand, question (L2) prompts us to contemplate the meaning of Lewis counterfactuals. As mentioned earlier, the truth-conditions of Lewis counterfactuals (and variably strict conditionals) are typically explained in terms of similarity between worlds: a counterfactual $A \Box\!\!\rightarrow B$ is true at a world w whenever B is true in the most similar A-worlds (i.e. worlds where A is true) to w. However, this truth conditional account, as we have seen, has been challenged from various perspectives, and alternative semantic accounts of counterfactuals have been proposed, such as truthmaker semantics (Briggs 2012; Fine 2012), causal modeling semantics (Galles and Pearl 1998; Pearl 1988), inquisitive semantics (Ciardelli, Zhang, and Champollion 2018), and team semantics (Barbero and Sandu 2020). Nonetheless, the results presented in this thesis suggest the possibility of an alternative interpretation of the truth conditions of Lewis counterfactuals, which may not necessarily rely on the requirement for the consequent to be true in the most similar worlds where the antecedent is true. Proposing a different interpretation of Lewis counterfactuals serves as more than just a philosophical exercise: it aids in understanding the true scope of Lewis' account and the nature of the counterfactual conditional that Lewis actually accounts for.

The prevailing paradigm suggests that Lewis counterfactuals are true when the consequent is true in the most similar worlds where the antecedent is true. However, an alternative interpretation of the standard truth conditions raises doubts about whether Lewis' account is correct or complete. For instance, it may suggest that the standard account is partial or misleading, and that Lewis' "counterfactuals" might pertain to a different relation between two propositions that is not transparent in the classical formulation of their truth conditions. A preliminary idea in this direction can be traced back to van Fraassen (1974) who reveals *hidden variables* in Lewis' account of counterfactuals. And by making these variables explicit, a more precise interpretation of Lewis' truth conditions for counterfactuals can be achieved, implying that the "most-similar-worlds" analysis might be a surface phenomenon emerging from a deeper truth-conditional account. Van Fraassen draws an analogy with a similar case in physics, where a theory T describing and predicting the behavior of a particular system can be extended to another theory $\langle T, \lambda \rangle$ here λ represents a *hidden* variable of the original theory. The extended theory with hidden variables is successful if it can make the same (correct) predictions as the original theory T while providing a more complete interpretation of T. In this context, $\langle T, \lambda \rangle$ reveals a certain parameter that was implicit in the original theory, thereby offering greater explanatory power and a better interpretation of T. Van Fraassen suggests that a similar situation could arise for Lewis counterfactuals by constructing a new theory based on Stalnaker conditionals, which can account for Lewis counterfactuals and interpret their truth conditions differently. Although these observations may currently appear obscure, a more detailed argument, along with an overview of van Fraassen's (1974) work, will be presented later in this thesis. For now, it is sufficient to keep in mind that question (L2) is not merely a philosophical exercise: it seeks to explore the possibility of providing a more comprehensive account of the truth conditions of conditionals that can also explain Lewis truth conditions for counterfactuals. Consequently, the urgency of addressing and answering question (L2) lies in the potential for offering a superior alternative to Lewis' paradigm regarding the truth conditions of his counterfactuals.

0.2.3 Counterfactuals and Probability

Another unresolved problem within the standard account pertains to characterizing the probability of the proposition expressed by a Lewis counterfactual (or a variably strict conditional in general). To clarify, let us assume, following the standard account, that the proposition expressed by a sentence A is the set of possible worlds at which A is true. Then, the probability of the proposition expressed by a sentence A, or equivalently, the probability that A is true, is the cumulative sum of the weights of the worlds where A is true (Lewis 1976). Thus, given a model with an underlying set of possible worlds W and a probability distribution P over W such that $\sum_{w \in W} P(w) = 1$, the probability that a sentence A is true is represented as $\sum_{w \models A} P(w)$, i.e. the cumulative sum of the probabilities of the worlds where A holds[4]. A characterization of the probability of a Stalnaker conditional, corresponding to variably strict conditional in the logic **VCS** in (Lewis 1973b) (or equivalently **C2** in Lewis 1971), has been provided by Lewis (1976). Specifically, Lewis proved that, given a suitable probability distribution P, the probability of a Stalnaker conditional, $A > B$ can be characterized as the probability of the consequent B, *imaged* on the antecedent A, i.e. $P(A > B) = P_A(B)$. The basic idea behind this characterization, is that, in the context of a probability distribution P on the underlying set of worlds W, imaging on A involves updating P to a new distribution P_A obtained by transferring the original mass of each not-A-world w ($P(w)$) to its closest A-world. The uniqueness of the closest world is ensured by the constraints on the Stalnakerian models for conditionals (a more detailed review of this characterization and the imaging procedure will be provided later in the thesis). On an intuitive level, what this characterization result tells us is that asking how probable it is that a Stalnaker conditional $A > B$ is true amounts to asking how probable it is that B is true after imaging that A holds. In other words, the probability that $A > B$ is true coincides with the probability that B is true, under the assumption that A holds in the imaged scenario. However, a similar characterization of the probability of a Lewis counterfactual (corresponding to the conditional in the logic **VC**, Lewis 1973b) is still lacking. On an intuitive level, such a characterization result would establish how to interpret the probability that a

[4] We have been intentionally sloppy at this stage. Later in the thesis, along the line of Lewis (1976), we will review in a more systematic way how to define a probability distribution over the worlds in a model that suitably extends to a probability function for all the propositions in that model.

counterfactual $A \mathrel{\Box\!\!\rightarrow} B$ is true, i.e. what it means to ask how probable it is that "If A were the case, then B would be the case", $P(A \mathrel{\Box\!\!\rightarrow} B) =?$. For instance, consider the counterfactual statement "If I hadn't received the vaccination, I would have caught the flu". While I may be quite certain that I would have caught the flu, if I hadn't received the vaccination, I cannot assign an absolute certainty to this counterfactual since there are various factors at play (e.g., immunity, contact with contagious individuals, etc.). The question then arises: how can we quantify the degree of confidence or probability associated with this counterfactual statement? That is, what is the probability assigned to this counterfactual? Does it coincide with the conditional probability of the consequent given the antecedent? Answering these questions and providing an interpretation of uncertainty quantification for Lewis counterfactuals are central objectives of this thesis.

At this point, some clarifications are necessary. Initially, we stated that, to the best of our knowledge, a characterization of the probability of Lewis counterfactuals (i.e. a conditional in the logic **VC**) is still lacking. However, upon closer examination, this claim may sound excessively strong and insufficiently justified. In fact, there are several different methods available for assigning a probability to a counterfactual conditional, and these methods often equate the probability of a counterfactual $A \mathrel{\Box\!\!\rightarrow} B$ with the probability of the consequent B under the counterfactual assumption that A holds, $P(A \mathrel{\Box\!\!\rightarrow} B) = P(B \; counterfactually \; given \; A)$. Different ways of specifying "*counterfactually given*" results in diverse approaches to defining or characterizing the probability of a counterfactual. For instance, Skyrms (1980a) have proposed understanding "*counterfactually given*" in terms of propensities, suggesting a conceptual modification of Adams-Edgington's (Adams 1965; Edgington 2008) prior epistemic probabilities view. Skyrms suggested to evaluate the degree of assertability of a counterfactual in terms of prior propensities instead of epistemic probabilities. This conceptual shift entails that the evaluation of a counterfactual is not reliant on *subjective* credence, as epistemic probabilities are, but rather on *objective* prior propensities represented by conditional chances. Skyrms' proposal suggests assessing the assertability of a counterfactual $A \mathrel{\Box\!\!\rightarrow} B$ based on the prior objective propensity of B given A, rather than our subjective belief in the prior conditional probability of B given A, as maintained by the prior epistemic probability view. Schulz (2017) further developed Skyrms' idea and presented a formal approach to specifying conditional propensity in terms of counterfactual conditional chances. Schulz's also suggests that, under suitable assumptions, its account of the probability of a counterfactual can be characterized in terms of *generalized imaging*, which was originally introduced by Gärdenfors (1982). Generalized imaging extends Lewis' imaging procedures to cases where the uniqueness assumption of the closest world is dropped. The fundamental idea behind this general rule is that, given a probability distribution P on the underlying set of worlds, imaging on A amounts to updating P to a new probability P^A obtained by redistributing the original mass of each not-A-world w among its closest not-A-worlds, such that each of these closest worlds receives a certain fraction of the original mass $P(w)$. To be precise, the generalized imaging encompasses a class of imaging rules that depend on how the original mass is redistributed. Consequently, Schulz (2017) suggests that the probability of a counterfactual $A \mathbin{\Box\!\!\rightarrow} B$ can be equated with the probability of B *generally imaged* on A, i.e. $P(A \mathbin{\Box\!\!\rightarrow} B) = P^A(B)$. This equation is further supported by Günther (2022), who reviews various ways of specifying generalized imaging and analyzes its connections with conditional probability and counterfactuals. Other accounts of the probability of counterfactuals have been developed by Leitgeb (2011), who introduced a probabilistic semantics for counterfactuals, and more recently by Santorio (forthcoming), who proposes understanding the probability of a counterfactual in terms of particular counterfactual conditional chances. Additionally, Pearl's (2000) account can be seen as equating the probability of a counterfactual with a certain counterfactual probability. Specifically, Pearl (2000) suggests that the probability of a counterfactual $A \mathbin{\Box\!\!\rightarrow} B$ corresponds to the probability of the consequent B, under an intervention $do(A)$ that forces the antecedent to be true, i.e. $P(A \mathbin{\Box\!\!\rightarrow} B) = P(B \mid do(A))$.

While all these accounts can be regarded as different specifications of the equation $P(A \mathbin{\Box\!\!\rightarrow} B) = P(B\ counterfactually\ given\ A)$, we will argue later in the thesis that none of these accounts successfully characterizes the probability of a Lewis counterfactual. In particular, the equation $P(A \mathbin{\Box\!\!\rightarrow} B) = P(B\ counterfactually\ given\ A)$ does not hold when the left-hand side term is understood as the probability of (the proposition expressed by) a Lewis counterfactual. In the light of this observation, it is now more accurate to claim that a characterization of the probability of a Lewis counterfactual is still lacking. Therefore, to address all these issues, we pose the following question regarding the probability of Lewis counterfactuals:

Question - Probability

(P1) How can we characterize the probability of (the proposition expressed by) a Lewis counterfactual? In other words, how can we interpret the probability that a Lewis counterfactual is true?

A comprehensive answer to the this question will be developed in Chapter 4 of the thesis. This answer will arise from a technical result that establishes a deep connection between Lewis counterfactuals and Dempster-Shafer theory of evidence (Dempster 1968; Shafer 1976). This result can also be seen as a generalization of Lewis' (1976) characterization of the probability of a Stalnaker conditional, with Lewis' result emerging as a specific instance of this more general characterization theorem.

Chapter 5 will also address a separate technical problem related to the probability of counterfactuals. As mentioned earlier, causal modeling semantics for counterfactuals (Galles and Pearl 1998; Pearl 2000) offers better insights into the probability of a counterfactual. Within a causal model, the probability of a counterfactual $A \mathrel{\Box\!\!\rightarrow} B$ can be interpreted as the probability that the consequent is true after an intervention that forces A to be true, i.e. $P(A \mathrel{\Box\!\!\rightarrow} A) = P(B \mid do(A))$[5]. However, causal modeling semantics presents some limitations when evaluating complex counterfactuals, particularly those with disjunctive antecedents such as $(A \vee B \mathrel{\Box\!\!\rightarrow} C)$ (e.g. "If it were raining or snowing, I wouldn't have gone to swim"), for standard causal modeling semantics cannot assign a truth-value or probability to such counterfactuals. Let us focus on this issue for a moment. First, causal modeling semantics, since the work of Galles and Pearl (1998), has emerged as a an alternative semantic account for counterfactuals. Specifically, a Pearl's (2000) model for counterfactuals consists in variables related to each other via a Bayesian network. Within a causal model, conditional dependencies among the variables can be analyzed. For instance, we can check in a causal model what happens to a certain variable V_2 *if* another variable V_1 is manipulated in a certain way. The concept of intervention amounts exactly to a certain manipulation of the variables in a model: we could intervene on a variable V and force it to assume a certain value v. The impact of this manipulation will propagate through the model potentially affecting other variables too. Galles and Pearl (1998) showed how this procedure can be employed to provide a semantic account of counterfactual conditionals: a counterfactual of the form $V_1 = 1 \mathrel{\Box\!\!\rightarrow} V_2 = 1$ is true in a causal model containing V_1 and V_2 if the intervention setting $V_1 = 1$ implies that also the variable V_2 takes value 1. We will review the causal modeling semantic framework in more details later in the thesis. For now, it is sufficient to highlight that this same semantic idea can be employed to assign a probability to a counterfactual: the probability of $V_1 = 1 \mathrel{\Box\!\!\rightarrow} V_2 = 1$ in a causal model amounts to the probability that $V_2 = 1$ holds after having performed the intervention $do(V_1 = 1)$. However, on the semantic level, Briggs (2012)

[5]It is important to point out that counterfactual conditionals within causal modeling semantics are not equivalent to Lewis counterfactuals: they have distinct truth-conditions and their logical framework differs from Lewis' counterfactual logic **VC** (e.g. see Halpern 2013). However, Galles and Pearl (1998) and Pearl (2017) demonstrated some interesting correspondences between the truth-conditions of Lewis counterfactuals and those of counterfactual conditionals interpreted within causal modeling semantics.

has showed that causal modeling semantics can assign a truth value only to a small class of counterfactuals. More precisely, counterfactuals with disjunctive antecedents such as $(V_1 = 1 \vee V_2 = 1) \mathrel{\Box\!\!\rightarrow} V_3 = 1$ lack a truth value within causal modeling semantics. This limitation is due to the fact that a *disjunctive intervention* is not defined: what does it mean to intervene on a causal model to set $V_1 = 1$ *or* $V_2 = 1$? Should we intervene on both V_1 and V_2? Briggs (2012) proposed a solution to this limited expressive power of causal modeling semantics by employing resources from truthmaker semantics (Fine 2017). However, we argue that a similar problem emerges at the probabilistic level: the probability of counterfactuals of the form $(V_1 = 1 \vee V_2 = 1) \mathrel{\Box\!\!\rightarrow} V_3 = 1$ cannot be computed within a causal model since disjunctive interventions such as $do(V_1 = 1 \vee V_2 = 1)$ are not defined, and Briggs' solution cannot be easily applied to the probabilistic case too. Hence, these considerations lead to another natural open problem:

Question - Probability

(P2) How can we characterize the probability of (the proposition expressed by) a counterfactual with disjunctive antecedent within causal modeling semantics? In other words, how can we interpret the probability that a counterfactual of the form $(A \vee B) \mathrel{\Box\!\!\rightarrow} C$ is true in a causal model?

0.3 Overview and Structure of The Thesis

In Chapter 1, we will develop a systematic algebraic treatment of variably strict conditional logics aiming to address questions (A1)-(A1a)-(A1b). Inspired by Nute (1975), we introduce algebraic structures, called *conditional algebras*, to investigate variably strict conditional logics. This approach reveals significant connections between variably strict conditional logics and modal logics. Previous works by Chellas (1975) and Segerberg (1989) have already recognized and explored the fruitful connection between conditional connectives and relative-necessity modal operators. They have shown how formal techniques from modal logics can be employed to prove some meta-logical facts, such as decidability, in conditional logics. Furthermore, building upon the research of Williamson (2010) and Lewis (1973b), Weiss (2019) has demonstrated how basic modal logic systems can be embedded into appropriate systems of variably strict conditional logics. In our study, we will deepen this connection between modal and variably strict conditional logics by differentiating between the *local* and *global* logical consequences associated with each variably strict conditional logic, in a way that mirrors the differentiation between global and local consequence relations in modal logics (see Blackburn, de Rijke, and Venema 2001; Wen 2021). Within this framework, we will demonstrate that the global-consequence counterpart of each variably strict conditional logic is algebraizable with respect to the corresponding class of conditional algebras. On the other hand, conditional algebras serve as an algebraic semantics for the local-consequence counterpart, which, on the other hand, is not algebraizable. Additionally, we initiate a structural investigation of conditional algebras by analyzing their deductive filters. The results of Chapter 1 can be schematically summarized as follows:

Summary of Chapter 1:

1. Introduction of *conditional algebras*, which amount to Boolean algebras equipped with a binary conditional operator;

2. Logical investigation of Lewis' variably strict conditional logics by differentiating between *global* and *local* consequence relations associated with each of those logics. We establish connections between these two types of consequence relations, similar to the relationship between local and global consequences in modal logic.

3. Demonstration that conditional algebras provide an equivalent algebraic semantics for the global consequence relations associated with Lewis' variably strict conditional logics, while the local consequence counterpart cannot be algebrized.

Chapter 2 of the thesis focuses on an innovative algebraic approach to counterfactuals, employing the framework of Boolean algebras of conditionals (BACs) introduced by Flaminio, Godo, and Hosni (2020). Initially, we introduce the BACs and establish a connection between the findings of Flaminio, Godo, and Hosni (2020) and the relevant philosophical literature on conditional probability and indicative conditionals. Our aim is to propose that the BACs represent an initial step towards "reconciling" the *suppositional* and *propositional* (or truth-conditional) perspectives on conditionals. The suppositional view, advocated for instance by Edgington (2008), suggests that conditional statements of the form "If A, then B" don't express a proposition or possess truth-conditions in the traditional sense. Instead, they express the conditional assertion of the consequent B under the supposition of the antecedent A. According to suppositional theorists, the evaluation of conditional assertions relies on conditional probability: evaluating "if A, then B" depends on the conditional probability of B *given* A, and the statement is considered truly assertible when the B is sufficiently plausible *given* A.

On the other hand, the propositional view, supported by Lewis (1976), Stalnaker (1968), Bennett (2003), posits that a conditional statement does express a proposition and possesses truth-conditions similar to other linguistic connectives. According to the propositional view, "if A, then B" is evaluated as *true* when certain conditions hold in the world of evaluation.

The incompatibility between these two perspectives is reinforced by notable results on conditionals and conditional probability known as *triviality results*. Lewis (1976) demonstrated that, under suitable and intuitive assumptions about the behavior of the probability function, the probability of the proposition expressed by a conditional "if A, then B" collapses to the probability of the consequent B. Before discussing this result, some preliminary clarifications are required. When referring to the standard truth-conditional account, we mean the view that the proposition expressed by a sentence corresponds to the set of possible worlds at which that sentence is true. For example, the proposition expressed by $A \wedge B$ in a model $\mathcal{M}$ is the intersection of the set of worlds where A is true and the set of worlds where B is true. Therefore, when dealing with a conditional binary connective $A > B$, specifying its truth-conditions at a world becomes necessary in order to determine the set of worlds defining the proposition expressed by $A > B$. Within this framework, given a probability distribution P over the underlying set of worlds, the probability of the proposition expressed by a sentence A, denoted $P(A)$, is the cumulative sum of the probabilities of the worlds in the proposition expressed by A, i.e. $P(A) = \sum_{w \models A} P(w)$ (see Section 0.2.3). Based on this background, Lewis' triviality result implies that, under suitable assumptions, the probability of the proposition expressed by $A > B$, where $>$ is a binary conditional connective, collapses to the unconditional probability of the consequent B, i.e. $P(A > B) = P(B)$, which is an undesired consequence. These triviality results have been further generalized. For instance, Hájek (1989) demonstrated that, under very minimal assumptions, it is not possible to find any binary connective $>$ within the standard truth-conditional account such that the probability of the proposition expressed by $A > B$ coincides with the conditional probability of B on A, without leading to a contradiction. Thus, we find ourselves at a crossroad: either we accept the standard truth-conditional account and relinquish the idea that the evaluation of a conditional is tied to the corresponding conditional probability, or we reject the truth-conditional account and adopt the suppositional perspective that a conditional statement "if A, then B" is considered truly assertible when the consequent B is sufficiently plausible given A. Several attempts have been made to restore the equation between the conditional probability and the probability of (the proposition expressed by) the the corresponding conditional, $P(A > B) = P(B \mid A)$ (e.g. Stalnaker 1970; van Fraassen 1976), and the BACs framework indeed falls into this category.

We will delve into a detailed examination of the BACs framework in Chapter 2 and its logical property. We will also briefly review some results connecting BACs and probability placing them in the discussion over the triviality results and the probability of conditionals. In this context, we show how the BACs framework effectively mitigates a problematic source of the triviality result, which was identified by Hájek (1989), thereby blocking the triggers leading to triviality results and reestablishing the connection between conditionals and conditional probabilities. In Chapter 2, we will mainly focus on the connections between the BACs framework and logics of conditionals. The main outcome of this chapter is the discovery that, within the BACs framework, Lewis first-degree counterfactuals (i.e. counterfactuals that do not contain any other nested counterfactual in either the antecedent or the consequent) can be defined using a normal modal operator and a conditional connective, which can be identified with Adams indicative conditional. Consequently, this addresses question (L1) by providing an answer to questions (L1b) and (L1c). This characterization of Lewis counterfactuals serves as an evidence supporting the idea that the assessment of a counterfactual (or subjunctive) conditional depends on the corresponding probabilistic/indicative conditional, thereby rekindling the idea proposed by Adams (1965) and Edgington (2008).

In the context of (L2), we will discuss our findings and argue that our results reveal counterfactuals as modalities of their corresponding indicatives, subtly aligning with Adams-Edgington's view on counterfactuals. Specifically, according to the prior epistemic probabilities view, a counterfactual can be seen as a *past* indicative, in the sense that it aligns to the past behavior (i.e. past modality) of its corresponding indicative conditional. Similarly, the hypothetical prior epistemic probability view suggests that that a counterfactual can be understood as a *hypothetical (past)* indicative, aligning to a certain hypothetical past behavior of the corresponding indicative. Therefore, our results provide a logical evidence supporting the idea that Lewis counterfactuals can be understood as modalities of their corresponding probabilistic/indicative conditionals. It is worth noting that the idea of counterfactual conditionals expressing a modality is not entirely new. Starr (2013) discusses two perspectives from linguistics and formal semantics that support this idea. The first is the *past as remote modality view* (supported, for example, by Iatridou 2000; Schulz 2007) which suggests that the past tense in subjunctive conditionals acts as a modal function, indicating that the antecedent of the subjunctive conditional expresses a possibility not assumed in the current context of discourse. The alternative view (supported, for instance, by Ippolito 2006; Khoo 2015) holds that the past tense in subjunctive conditionals conveys a modality expressing what was necessary in the past. However, these ideas are primarily based on linguistic observations and have not been extensively formalized. Additionally, they are applied to subjunctive conditionals in general, rather than specifically addressing Lewis counterfactuals. Nevertheless, within the context of this thesis, they can be seen as an additional evidence, originating from a different field, supporting the idea that counterfactuals, as subjunctive conditionals, express a specific modality. This observation, combined with the preceding discussion, sheds light on the results presented in Chapter 2. We will argue that these results not only represent a technical advancement for the logic of conditionals but also serve as a small step towards implementing a reductionist approach to understanding the truth conditions of counterfactuals. Specifically, counterfactual conditionals may be reduced to a combination of a probabilistic conditional and a modal operator. The interpretation of this modal operator can vary, and we will explore different plausible options. The results in Chapter 2 can be schematically summarized as follows:

Summary of Chapter 2:

1. Examination of Boolean algebras of conditionals (BACs) introduced in Flaminio, Godo, and Hosni (2020) and exploration of their properties; philosophical observations connecting BACs with the classical philosophical debate on conditional probabilities vs probability of conditionals;

2. Introduction of a novel type of algebraic structures, which we call "Lewis algebras", which consist of BACs equipped with a specific normal modal operator;

3. Analysis of the structural properties of Lewis algebras and their dual structures using the Jónnson-Tarski duality theory for modal algebras; the dual structures of our Lewis algebras are a special class of Kripke frames that reflect Lewis sphere models for counterfactuals;

4. Proof of soundness and completeness results for a slightly stronger version of Lewis' logic of Counterfactuals **VC** (= **C1** in Lewis 1971) with respect to our Lewis algebras and their corresponding dual frames;

5. From a philosophical standpoint, discussion of the aforementioned technical results, arguing that they offer a new interpretation of Lewis counterfactuals based on a reductionist perspective, wherein a counterfactual can be characterized as a modality of the corresponding probabilistic/indicative conditional.

Chapter 3 builds upon the findings of Chapter 2 to specifically address questions (L1) and (L2) by providing an answer to questions (L1a) and (L1c). We extend the results obtained in the BACs framework concerning the characterization of counterfactuals to encompass the entire Lewis language, including embedded counterfactuals and the full class of variably strict conditionals. To achieve this, we introduce a new class of models, called *spherical Kripke models*, which allow us to define Lewis variably strict conditionals using a normal modal operator combined with Stalnaker conditionals. This definability is made possible by enriching Stalnaker conditional language, which contains the conditional operator $>$, with a modal operator $\Box$ that can be interpreted within a spherical Kripke model. In the context of this expanded language, Lewis counterfactuals/variably strict conditionals can be characterized as formulas of the form $\Box(A > B)$, i.e. $A \Box\!\!\rightarrow B \equiv \Box(A > B)$. We argue how this result can be regarded as a refinement and an extension of van Fraassen's (1974) characterization results for Lewis counterfactuals. While van Fraassen showed that a special family of Stalnakerian models can induce a Lewisian model for counterfactuals, we demonstrate how our spherical Kripke models encode both a Stalnakerian and a Lewisian model, and how the latter can be translated into an equivalent spherical Kripke model. Additionally, we axiomatize the resulting logics induced by our spherical Kripke models, revealing interesting features of these logics, which we refer to as **KV**-logics. These logics can be viewed as modal extensions of Stalnaker's conditional logic **VCS** and represent a new instance of *weak logics*, meaning they are not closed under uniform substitution. The notion of weak logics has been very recently introduced by Nakov and Quadrellaro (2022) to classify a particular class of logics used, for instance, to model epistemic scenarios (e.g. public announcement logics Holliday, Hoshi, and Icard 2013) and reasoning with questions (e.g. inquisitive logics Ciardelli 2022). The non-closure under uniform substitution makes **KV**-logics intriguing as they require a specialized algebraic treatment beyond the standard framework of algebraic logic. Although we do not extensively explore the algebraic perspective of these logics, it is noteworthy that they represent a novel and significant example within the class of weak logics. Furthermore, we demonstrate that each of Lewis' variably strict conditional logics can be embedded into its corresponding **KV**-logic.

The characterization of Lewis counterfactuals within the expanded language of **KV**-logics and the embeddability of variably strict conditional logics into **KV**-logics provide an opportunity for a philosophical investigation into question (L2). In Chapter 3, we evaluate the philosophical implications of these results by employing conceptual frameworks developed in existing literature.

First, we assess the results in relation to van Benthem's (2018) dichotomy of *explicit* vs *implicit* stances in logic. We argue that the relationship between variably strict conditional logics and **KV**-logics serves as further evidence for the existence of this dichotomy, aligning with van Fraassen's idea of Lewis' theory with hidden variables. According to van Benthem's terminology, implicit stances in logic involve modifying and enriching the meaning of logical constants, leading to new systems where the meaning of those constants deviated from standard classical logic. Intuitionistic logic, for instance, exemplifies an implicit stance by altering the classical meaning of negation. In contrast, explicit stances involve enriching the classical vocabulary with new constants to model new phenomena. Modal logic is a notable example, as it introduces new operators to classical logic to capture concepts like knowledge, beliefs, and alethic modalities. Sometimes, these stances represent two sides of the same coins. For instance, intuitionistic logic can be embedded (translated) into the modal logic **S4** (Gödel 1986; Kripke 1965), thereby making what was *implicit* in intuitionistic logic explicit in the new expanded vocabulary of **S4**.

This offers a new interpretation of intuitionistic logic as well. Drawing from this perspective, we argue that Lewis' variably strict conditional logics can be seen as an *implicit* stance in logic, while our **KV**-logics as their *explicit* counterpart, mirroring the implicit-explicit relation between intuitionistic logic and modal logic **S4**. We suggest that van Fraassen (1974) already captured the intuition behind this implicit-explicit distinction by proposing an extended theory (explicit stance) that made Lewis' hidden variables (implicit stance) explicit.

Furthermore, we contend that the embeddability of variably strict conditional logics into **KV**-logics and the characterization of Lewis counterfactuals in terms of Stalnaker conditional and a modal operator offer a new interpretation of Lewis' logics, addressing question (L2). Specifically, the fact that Lewis counterfactuals can be seen as Stalnaker conditionals preceded by a modal operator allows for a fresh understanding of their truth conditions. Within spherical Kripke models, a counterfactual $A \mathrel{\Box\!\!\rightarrow} B$ is deemed true when the truth-conditions of $\Box(A > B)$ are satisfied, indicating that a certain conditional $A > B$ must be *necessary*. This raises the question of how to interpret this necessity, whether as alethic, epistemic, or deontic. We refrain from committing to a particular interpretation and instead propose several interpretations that seem plausible in light of existing literature. For example, we argue that interpreting this necessity as *epistemic* seems plausible, paralleling the epistemic interpretation of intuitionistic logic into modal logic **S4** as proposed by Kripke (1965). In this sense, a counterfactual, now denoted as $\Box(A > B)$, would be true when the corresponding conditional $A > B$ is *known*. Additionally, we also argue that an interpretation of the relevant modality in terms of *provability* is also plausible and supported by existing evidence in the literature (Gödel 1986; Weiss 2019). According to this interpretation, a counterfactual $\Box(A > B)$ would be true when the corresponding conditional $A > B$ is *provable*.

The result in Chapter 3 can be schematically summarized as follows:

Summary of Chapter 3:

1. Introduction of a novel kind of models called "spherical Kripke models", which extend Stalnakerian conditional models by incorporating an accessibility relation;

2. Introduction of a new class of Logics called "**KV**-logics", which are formulated in modal extension of Stalnaker's language; proof of soundness and completeness of **KV**-logics with respect to spherical Kripke models;

3. Characterization of Lewis counterfactual connective, within the framework of **KV**-logics, in terms of a Stalnaker conditional preceded by a modal operator; embeddability/translation results of Lewis' variably strict conditional logics into **KV**-logics;

4. Philosophical exploration of the implications of the embeddability/characterization results of Lewis counterfactuals within the context of the dichotomy of implicit-explicit stances in logic;

5. Proposal of a new interpretation of the truth-conditions of Lewis counterfactuals, linking them to special modalities of the corresponding Stalnaker conditionals.

Chapter 4 addresses question (P1) regarding the probability of counterfactuals. This chapter makes several key contributions, including the characterization of counterfactual probability. Additionally, we introduce the Dempster-Shafer (1968; 1976) theory of belief functions to the philosophical discussion and demonstrate its potential for providing simple solutions to significant philosophical problems.

The chapter begins by presenting the challenge of characterizing the probability of (the proposition expressed by) a Lewis counterfactual. We examine existing accounts, such as that proposed by (Günther 2022) and demonstrate its failure to capture the probability of Lewis counterfactuals. To address this issue, we introduce and review the framework of belief functions (Dempster 1968; Shafer 1976). Specifically, we delve into the relationship between belief functions and modal logic by reviewing prior research in the field (Hájek 1996; Harmanec, Klir, and Wang 1996; Resconi, Klir, and Clair 1992). These studies demonstrate how modal logic can be interpreted as the foundational logic for belief functions. In other words, each belief function can be derived from a specific Kripke frame equipped with a probability distribution over the underlying set of worlds. In a slogan, classical logic corresponds to classical probability theory as modal logic aligns with belief function.

The aforementioned results, in combination with our findings in Chapters 2 and 3 that demonstrate the characterization of a Lewis counterfactual in terms of a Stalnakerian/Adams-like conditional preceded by a modal operator, lead to the inference that the probability of a Lewis counterfactual can be characterized via a belief function. This belief function can be further characterized in terms of an *imaged*-belief function, which serves as a special updating rule generalizing Lewis' imaging to a non-Bayesian context. In particular, we show how this kind of belief updating differs from Gärdenfor's (1982) generalized imaging in that the latter *reallocates* the lost mass of a world among its nearest worlds (see Section 0.2.3). In contrast, our imaged-belief function *transfers* the entire lost mass of a world to the set of its closest worlds, thus inducing a belief function. We contend that this result can be viewed as a formal advancement of an observation by Dubois and Prade (1994), who suggested that Lewis' models of counterfactuals give rise to an *imaged*-belief function within the context of Dempster-Shafer. Indeed, within this framework, we argue that Lewis' logic of counterfactuals possesses a probabilistic interpretation. Specifically, Lewis' axioms of **VC** (as well as other axioms in variably strict conditional logics) impose constraints on the imaged belief function induced by the corresponding **VC** Lewisian models. Furthermore, we propose that this interpretation also suggests that the modality $\Box$ involved in the characterization of Lewis counterfactuals, i.e. $A \Box\!\!\rightarrow B \equiv \Box(A > B)$ can be understood in terms of a *provability* modality. This notions aligns with Pearl's (1988) observation, which posits that the belief function of a formulas can be seen as a measure of the *degree of provability* of that formula. More schematically, the results of Chapter 4 can be summarized as follows:

Summary of Chapter 4:

1. Introduction of Dempster-Shafer's framework of belief function and review of the classical results connecting modal logic and belief functions;

2. Characterization of the probability of Lewis counterfactuals using belief functions; characterization of the belief function associated with a counterfactual as a special type of imaged belief;

3. Proposal of a probabilistic interpretation of variably strict conditional logics;

In Chapter 5, we focus specifically on addressing question (P2). Drawing on Briggs' work (2012), we present a procedure for assigning probabilities to counterfactuals of the form $(A \vee B) \Box\!\!\rightarrow C$ within Pearl's causal models. We begin by reviewing Pearl's causal modeling semantics for counterfactuals and its limitations, particularly its inability to assign truth values to counterfactuals with disjunctive antecedents like $(A \vee B) \mathbin{\Box\!\!\rightarrow} C$. We then examine Briggs's solution to this expressive power issue using truthmaker semantics. Briggs (2012) demonstrated how interventions on a causal model can be interpreted as *truthmakers* for the corresponding formulas. For example, an intervention like $do(X = 1)$, which sets the variable X to the value 1, can be seen as a truthmaker for the statement $X = 1$. Based on this idea, Briggs proposes that a counterfactual of the form $(A \vee B) \mathbin{\Box\!\!\rightarrow} C$ is true in a causal model if the consequent C is made true by all the truthmakers (i.e. interventions) that verify the antecedent $A \vee B$.

Similarly, we aim to extend this concept to the probabilistic framework. We employ a similarity measure among causal models introduced by Eva, Stern, and Hartmann (2019), which is based on the number of counterfactual dependencies. According to this measure, given a model $\mathcal{M}$, another model $\mathcal{M}_1$ is more similar to $\mathcal{M}$ than a model $\mathcal{M}_2$, if $\mathcal{M}_1$ agrees with $\mathcal{M}$ on a greater number of counterfactual dependencies among their variables. Consequently, we propose that the probability of a counterfactual $(A \vee B) \mathbin{\Box\!\!\rightarrow} C$ can be computed as a weighted average of the probability of C with the respect to the truthmakers (i.e. intervened models) of the antecedents. The weight assigned to each truthmaker is proportional to its similarity to the original model. Furthermore, we compare the predictions generated by our procedure with those of Lewis' imaging. Pearl (2017) showed that assigning a probability to a counterfactual in a causal model can be interpreted as a Lewis' imaging rule. However, we contend that our procedure outperforms Lewis' imaging, especially when applied to counterfactuals with disjunctive antecedents. In fact, Lewis' imaging violates a plausible *convexity principle* that we propose as a reasonable constraint for assigning probabilities to a counterfactual.

More schematically, the results of Chapter 5 can be summarized as follows:

Summary of Chapter 5:

1. Introduction and review of the problem of limited expressive power for causal modeling semantics and discussion of Briggs' (2012) solution;

2. Demonstration of how the limited expressive power of causal modeling semantics affects the assignment of probabilities to counterfactuals within a causal model;

3. Proposal of a new generalized procedure to compute the probability of counterfactuals in the context of a causal model obtained by combining Briggs' (2012) solution with the findings by Eva, Stern, and Hartmann (2019); characterization of the probability of a counterfactual with disjunctive antecedents in a causal model

4. Proposal of a convexity constraint as a reasonable requirement for assigning probabilities to counterfactuals in a causal model; comparison of our findings with Lewis' imaging procedure.

0.4 Preliminaries

In this section, we provide the basic technical background required for the results proved in the following chapters. Specifically, we review Lewis' systems of variably strict conditional logics and their associated semantics. All the definitions and results that we will introduce are taken from or are a direct extension of the results in (Lewis 1973b) and Lewis (1971). In what follows, since the notation may look heavy, in order to facilitate the reading, we put into a frame the information that is relevant for clarifying the notation.

0.4.1 Syntax and Semantics

Let us start with introducing our basic language. Let *Var* be an enumerable set of propositional variables; we use lower-case Latin letters $p, q, r, \ldots$ to denote variables in *Var*. Let $\mathcal{L}$ be a standard classical logical language in the signature $\neg, \vee, \wedge$ over *Var*, having round brackets as auxiliary symbols. We take the liberty to omit brackets when they appear in the outermost position and whenever the scope of a connective is clear from the context. We use lower-case Greek letters $\varphi, \psi, \delta, \ldots$ to indicate formulas in $\mathcal{L}$. The material implication $\supset$ is defined as usual: $\varphi \supset \psi := \neg\varphi \vee \psi$, as well as the bi-conditional $\leftrightarrow$: $\varphi \leftrightarrow \psi := (\varphi \supset \psi) \wedge (\psi \supset \varphi)$. Furthermore, we can also define the *truth* as $\top := p \vee \neg p$

Definition 0.1 (Lewis Language)**.** *$\mathcal{L}_{\Box\!\!\rightarrow}$ amounts to Lewis (1971; 1973) language for variably strict conditional logics. $\mathcal{L}_{\Box\!\!\rightarrow}$ consists in expanding $\mathcal{L}$ with the binary connective $\Box\!\!\rightarrow$, which stands for Lewis variably strict conditional. Formulas of $\mathcal{L}_{\Box\!\!\rightarrow}$ are defined as follows:*

- *if φ is a formula of $\mathcal{L}$, then φ is a formula of $\mathcal{L}_{\Box\!\!\rightarrow}$ too;*
- *if φ and ψ are formulas of $\mathcal{L}_{\Box\!\!\rightarrow}$ then so is $(\varphi \Box\!\!\rightarrow \psi)$;*
- *if φ, ψ are formulas of $\mathcal{L}_{\Box\!\!\rightarrow}$, then so are $\neg\varphi$, $\varphi \vee \psi$, and $\varphi \wedge \psi$*
- *nothing else is a formula of $\mathcal{L}_{\Box\!\!\rightarrow}$.*

$For_{\mathcal{L}_{\Box\!\!\rightarrow}}$ denotes the set of formulas of $\mathcal{L}_{\Box\!\!\rightarrow}$. Several non-primitive connectives are defined in $\mathcal{L}_{\Box\!\!\rightarrow}$:

- *the might counterfactual $(\varphi \Diamond\!\!\rightarrow \psi) := \neg(\varphi \Box\!\!\rightarrow \neg\psi)$;*
- *a box operator $\boxdot\varphi := \neg\varphi \Box\!\!\rightarrow \varphi$;*
- *a diamond operator, $\Diamond\varphi := \neg \boxdot \neg\varphi$;*
- *a comparative plausibility operator $\varphi \preccurlyeq \psi := ((\varphi \vee \psi) \Diamond\!\!\rightarrow (\varphi \vee \psi)) \supset ((\varphi \vee \psi) \Diamond\!\!\rightarrow \varphi)$, where $\varphi \preccurlyeq \psi$ can be read as "φ is at least as plausible as ψ".*

We will employ two different, but equivalent, interpretation structures for $\mathcal{L}_{\Box\!\!\rightarrow}$ introduced by Lewis (1971) and called "α-models" and "β-models". We stick to a different, yet more transparent, terminology to indicate these interpretation structures.

Definition 0.2. *A* spherical Lewisian model *is a tuple $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$ where:*

- *W is a non-empty set (of possible worlds);*
- *$\mathcal{S}$ is a function $\mathcal{S} : W \rightarrow \wp(\wp(W)) \setminus \emptyset$ such that for all $w \in W$, the following conditions hold:*
  - ***Nestedness**: for all $S, T \in \mathcal{S}(w)$, either $S \subseteq T$ or $T \subseteq S$;*
- *$\vDash$ is a valuation relation $\vDash \subseteq W \times Var$ that is extended to compound formulas of $\mathcal{L}_{\Box\!\!\rightarrow}$ as follows:*

$$
\begin{array}{lcl}
w \vDash \neg\varphi & \Leftrightarrow & w \nvDash \varphi \\
w \vDash \varphi \wedge \psi & \Leftrightarrow & w \vDash \varphi \text{ and } w \vDash \psi \\
w \vDash \varphi \vee \psi & \Leftrightarrow & w \vDash \varphi \text{ or } w \vDash \psi \\
w \vDash \varphi \Box\!\!\rightarrow \psi & \Leftrightarrow & \text{either for all } v \text{ such that } v \in \bigcup \mathcal{S}(w), v \nvDash \varphi, \text{or} \\
& & \text{there is a } S \in \mathcal{S}(w) \text{ and a } v \in S \text{ such that } v \vDash \varphi \\
& & \text{and for all } u \in S,\ u \vDash \varphi \supset \psi
\end{array}
$$

$\mathcal{M} \vDash \varphi$ means that for all $w \in W$, $w \vDash \varphi$. Moreover we define the proposition expressed by a formula φ as follows:

$$[\varphi] = \{w \in W \mid w \vDash \varphi\}$$

Thus, we can rewrite the clause for counterfactuals as follows:

$$w \vDash \varphi \,\Box\!\!\rightarrow \psi \quad \Leftrightarrow \quad \text{either } \bigcup \mathcal{S}(w) \cap [\varphi] = \emptyset \text{ or there is a } S \in \mathcal{S}(w) \text{ such that } S \cap [\varphi] \neq \emptyset \text{ and } (S \cap [\varphi]) \subseteq [\psi]$$

Moreover, we assume that $\mathcal{S}$ satisfies the following condition:

- ***Limit Assumption**: for all $\varphi \in For_{\mathcal{L}_{\Box\rightarrow}}$, if $[\varphi] \cap \bigcup \mathcal{S}(w) \neq \emptyset$, then there is a $S \in \mathcal{S}(w)$ such that $S \cap [\varphi] \neq \emptyset$ and for all $T \in \mathcal{S}(w)$, if $T \cap [\varphi] \neq \emptyset$, then $S \subseteq T$; such S may be denoted as the minimal φ-permitting sphere in $\mathcal{S}(w)$.*

- *for formulas $\varphi \in For_{\mathcal{L}_{\Box\rightarrow}}$ and for $w \in W$, we use $min^{\varphi}_{\subseteq}(\mathcal{S}(w))$ to denote the minimal φ-permitting sphere in $\mathcal{S}(w)$, if any; otherwise $min^{\varphi}_{\subseteq}(\mathcal{S}(w))$ would denote the empty-set.*
- *$[\varphi] \cap min^{\varphi}_{\subseteq}(\mathcal{S}(w))$ may be interpreted as the set of the closest φ-world to w.*

Remark 0.1. A spherical Lewisian model corresponds to a β-model for the logic **C0** defined by Lewis (1971) with the additional limit assumption. As Lewis (1973b, p.121) points out, adding the limit assumption is immaterial as it doesn't affect the resulting (finitary) logic associated with a certain class of spherical Lewisian models. As it will be clearer later, we have decided to opt for the limit assumption for the sake of simplicity in order to have a straightforward and neat correspondence between spherical and functional models.

In order to simplify the semantic conditions of counterfactual conditionals, we will show the following result:

Lemma 0.1. *For any spherical Lewisian model $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$, for all $w \in W$, for all $\varphi, \psi \in For_{\mathcal{L}_{\Box\rightarrow}}$, the following holds:*

$$w \vDash \varphi \,\Box\!\!\rightarrow \psi \Leftrightarrow ([\varphi] \cap min^{\varphi}_{\subseteq}(\mathcal{S}(w))) \subseteq [\psi]$$

Proof.

($\Rightarrow$) By contrapositio, assume that $[\varphi] \cap \bigcup \mathcal{S}(w) \neq \emptyset$ and $min^{\varphi}_{\subseteq}(\mathcal{S}(w))) \nsubseteq [\psi]$. This means that there is a $v \in W$ such that $v \in min^{\varphi}_{\subseteq}(\mathcal{S}(w)))$ but $v \notin [\psi]$. Moreover, since $\mathcal{S}(w)$ is totally ordered by set inclusion, we have that $v \in S$ for all $S \in \mathcal{S}(w)$ such that $S \cap [\psi] \neq \emptyset$. Therefore, for all $S \in \mathcal{S}(w)$, if $S \cap [\varphi] \neq \emptyset$ then $S \cap [\varphi] \cap [\neg\psi] \neq \emptyset$. And so $s \nvDash \varphi \mathrel{\Box\!\!\rightarrow} \psi$.

($\Leftarrow$) This direction is straightforward since either $min^{\varphi}_{\subseteq}(\mathcal{S}(w))) = \emptyset$ or, by the limit assumption, $min^{\varphi}_{\subseteq}(\mathcal{S}(w)))$ exists and $min^{\varphi}_{\subseteq}(\mathcal{S}(w))) \in \mathcal{S}(w)$.

□

Lewis (1973b) introduced several classes of spherical Lewisian models[6] and provided the characteristic axioms for each of those classes. These axioms, and the induced logics, are formulated in a different language that doesn't include the conditional connective $\mathrel{\Box\!\!\rightarrow}$ as primitive; however, Lewis (1973b, pp 132-133) provides useful insights on how to formulate all its variably strict conditional logics in the language $\mathcal{L}_{\Box\!\rightarrow}$. In the following Table 1, we summarize all the class of spherical Lewisian models considered by Lewis (1973b) and their characteristic axioms. Each class is identified by a specific condition on the system of spheres, and this condition is characterized by a corresponding axiom.

[6] To be precise, the spherical Lewisian models we introduced here are not exactly those defined by Lewis (1973b). We stick to the spherical Lewisian models as defined in (Lewis 1971); however, all the structural conditions that Lewis (1973b) considers also suit our definition of spherical Lewisian models.

	Spherical Lewisian Models	
	Condition	Axiom (in $\mathcal{L}_{\Box\!\!\rightarrow}$)
(N) Normality	$\bigcup \mathcal{S}(w) \neq \emptyset$	$\boxdot\varphi \supset \diamonddot\varphi$
(T) Total Reflexivity	$w \in \bigcup \mathcal{S}(w)$	$\boxdot\varphi \supset \varphi$
(W) Weak Centering	$w \in \bigcap \mathcal{S}(w)$	$(\varphi \mathrel{\Box\!\!\rightarrow} \psi) \supset (\varphi \supset \psi)$
(C) Centering	$\{w\} \in \mathcal{S}(w)$	**Weak Centering** + $(\varphi \wedge \psi) \supset (\varphi \mathrel{\Box\!\!\rightarrow} \psi)$
(S) Uniqueness	if $[\varphi] \cap \bigcup \mathcal{S}(w) \neq \emptyset$ then $\exists S \in \mathcal{S}(w) : \lvert[\varphi] \cap S\rvert = 1$	$(\varphi \mathrel{\Box\!\!\rightarrow} \psi) \vee (\varphi \mathrel{\Box\!\!\rightarrow} \neg\psi)$
(U-) Local Uniformity	for all $v \in \bigcup \mathcal{S}(w), \bigcup \mathcal{S}(w) = \bigcup \mathcal{S}(v)$	$\diamonddot\varphi \supset \boxdot\diamonddot\varphi$ + $\boxdot\varphi \supset \boxdot\boxdot\varphi$
(U) Uniformity	for all $w, v \in W, \bigcup \mathcal{S}(w) = \bigcup \mathcal{S}(v)$	$\diamonddot\varphi \supset \boxdot\diamonddot\varphi$ + $\boxdot\varphi \supset \boxdot\boxdot\varphi$
(A-) Local Absoluteness	for all $v \in \bigcup \mathcal{S}(w), \mathcal{S}(w) = \mathcal{S}(v)$	$(\varphi \preccurlyeq \psi) \supset \boxdot(\varphi \preccurlyeq \psi)$ + $(\varphi \prec \psi) \supset \boxdot(\varphi \preccurlyeq \psi)$
(A) Absoluteness	for all $w, v \in W, \mathcal{S}(w) = \mathcal{S}(v)$	$(\varphi \preccurlyeq \psi) \supset \boxdot(\varphi \preccurlyeq \psi)$ + $(\varphi \prec \psi) \supset \boxdot(\varphi \preccurlyeq \psi)$
(UT) Universality	$\bigcup \mathcal{S}(w) = W$	**Uniformity + Total Reflexivity**
(WA) Weak Triviality	$\mathcal{S}(w) = \{W\}$	**Weak Centering + Absoluteness**
(CA) Triviality	$\mathcal{S}(w) = \{\{w\}\}$	**Centering + Absoluteness**

Table 1: The table schematically summarizes the structural conditions over spherical Lewisian models and the corresponding characteristic axioms. The axioms schemata for spherical Lewisian models are meant to range over all formulas in $\mathcal{L}_{\Box\!\!\rightarrow}$; and the conditions are meant to range over all formulas and all possible worlds w.

We introduce the following notation to refer to classes of spherical Lewisian models and define logical consequence with respect to a certain class.

Notation 0.1.

- $\mathbf{V}$ *denotes the class of all spherical Lewisian models.*
- $\mathfrak{C}$ *indicates a condition or a family of conditions (possibly empty) among those in Table 1, i.e.* $\{\mathbf{N}, \mathbf{T}, \mathbf{W}, \mathbf{C}, \mathbf{S}, \mathbf{A}, \mathbf{U}, \}$.
- $\mathbf{V}\mathfrak{C}$ *denotes the class of spherical Lewisian models satisfying condition(s)* $\mathfrak{C}$
- *Logical consequence over a class of spherical Lewisian models is defined as follows: for* $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}_{\Box\!\rightarrow}}$,

$$\Gamma \models_{\mathbf{V}\mathfrak{C}} \varphi \quad \Leftrightarrow \quad \begin{array}{l} \text{for all spherical Lewisian models } \mathcal{M} \text{ satisfying conditions } \mathfrak{C}, \\ \text{for all } w \text{ in } \mathcal{M}, \text{ if } \mathcal{M}, w \vDash \gamma \text{ for all } \gamma \in \Gamma, \text{ then } \mathcal{M}, w \vDash \varphi \end{array}$$

For example, $\Gamma \models_{\mathbf{V}} \varphi$ means that φ is a logical consequence of Γ with respect to the class of all spherical Lewisian models; $\Gamma \models_{\mathbf{VC}} \varphi$ means that φ is a logical consequence of Γ with respect to the class of all spherical Lewisian models satisfying **Centering**.

It is straightforward to prove that each spherical Lewisian model satisfies a condition $\mathfrak{C}$ among those in Table 1 if and only if all the worlds in that model make the corresponding axiom $\mathfrak{C}$ true.

Lemma 0.2. *Let* $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$ *be a spherical Lewisian model. We have that*

$$\text{for all } \varphi \in For_{\mathcal{L}_{\Box\!\rightarrow}}, \mathcal{M} \vDash [\text{axiom } \mathfrak{C}] \Leftrightarrow \mathcal{M} \text{ satisfies condition } \mathfrak{C}$$

We will now introduce alternative model theoretic structures to interpret $\mathcal{L}_{\Box\!\rightarrow}$; indeed, for the central results of this thesis, it will be useful to go back and forth from all these different alternative models.

Definition 0.3. *A functional Lewisian model is a tuple of the form* $\langle W, f, \vDash \rangle$ *where*

- W *is a non-empty set (of possible worlds);*
- $f : For_{\mathcal{L}_{\Box\!\rightarrow}} \times W \rightarrow \wp(W)$ *is a function mapping a pair made of a formula and a world to a set of worlds;*

- $\models \subseteq W \times Var$ *is a valuation relation that is extended to compound formulas of* $\mathcal{L}_{\Box\!\!\rightarrow}$ *as follows:*

$$
\begin{array}{lcl}
w \models \neg\varphi & \Leftrightarrow & w \not\models \varphi \\
w \models \varphi \wedge \psi & \Leftrightarrow & w \models \varphi \text{ and } w \models \psi \\
w \models \varphi \vee \psi & \Leftrightarrow & w \models \varphi \text{ or } w \models \psi \\
w \models \varphi \Box\!\!\rightarrow \psi & \Leftrightarrow & \text{for all } v \in f(\varphi, w), v \models \psi
\end{array}
$$

$\mathcal{M} \models \varphi$ *means that for all* $w \in W$, $w \models \varphi$. *Moreover we define the proposition expressed by a formula* φ *as follows:*

$$[\varphi] = \{w \in W \mid w \models \varphi\}$$

Thus, we can rewrite the clause for counterfactuals as follows:

$$w \models \varphi \Box\!\!\rightarrow \psi \quad \Leftrightarrow \quad f(\varphi, w) \subseteq [\psi]$$

Moreover, we assume f *satisfies the following constraints: for all* $w \in W$, *for all* $\varphi, \psi \in For_{\mathcal{L}_{\Box\!\!\rightarrow}}$,

1. $f(\varphi, w) \subseteq [\varphi]$;
2. *if* $f(\varphi, w) \subseteq [\psi]$ *and* $f(\psi, w) \subseteq [\varphi]$, *then* $f(\varphi, w) = f(\psi, w)$;
3. *either* $f(\varphi \vee \psi, w) \subseteq [\varphi]$ *or* $f(\varphi \vee \psi, w) \subseteq [\psi]$ *or* $f(\varphi \vee \psi, w) = f(\varphi, w) \cup f(\psi, w)$;

Analogously to the case of spherical Lewisian models, we can identify specific classes of functional Lewis models characterized by a specific axiom, according to the following table:

	Functional Lewisian Models	
	Condition	Axiom (in $\mathcal{L}_{\boxright}$)
(N) Normality	for some $\varphi \in For_{\mathcal{L}_{\boxright}}$, $f(\varphi, w) \neq \emptyset$	$\boxdot\varphi \supset \diamonddot\varphi$
(T) Total Reflexivity	for some $\varphi \in For_{\mathcal{L}_{\boxright}}$, $w \in f(\varphi, w)$	$\boxdot\varphi \supset \varphi$
(W) Weak Centering	if $w \in [\varphi]$, then $w \in f(\varphi, w)$	$(\varphi \boxright \psi) \supset (\varphi \supset \psi)$
(C) Centering	if $w \in [\varphi]$, then $f(\varphi, w) = \{w\}$	**Weak Centering** + $(\varphi \wedge \psi) \supset (\varphi \boxright \psi)$
(S) Uniqueness	$\lvert f(\varphi, w)\rvert \leq 1$	$(\varphi \boxright \psi) \vee (\varphi \boxright \neg\psi)$
(U-) Local Uniformity	if $v \in \bigcup_{\varphi \in For_{\mathcal{L}_{\boxright}}} f(\varphi, w)$, then $\bigcup_{\varphi \in For_{\mathcal{L}_{\boxright}}} f(\varphi, w) = \bigcup_{\varphi \in For_{\mathcal{L}_{\boxright}}} f(\varphi, v)$	$\diamonddot\varphi \supset \boxdot\diamonddot\varphi$ + $\boxdot\varphi \supset \boxdot\boxdot\varphi$
(U) Uniformity	for all $w, v \in W$, $\bigcup_{\varphi \in For_{\mathcal{L}_{\boxright}}} f(\varphi, w) = \bigcup_{\varphi \in For_{\mathcal{L}_{\boxright}}} f(\varphi, v)$	$\diamonddot\varphi \supset \boxdot\diamonddot\varphi$ + $\boxdot\varphi \supset \boxdot\boxdot\varphi$
(A-) Local Absoluteness	if $v \in \bigcup_{\varphi \in For_{\mathcal{L}_{\boxright}}} f(\varphi, w)$, then for all $\varphi \in For_{\mathcal{L}_{\boxright}}$, $f(\varphi, w) = f(\varphi, v)$	$(\varphi \preccurlyeq \psi) \supset \boxdot(\varphi \preccurlyeq \psi)$ + $(\varphi \prec \psi) \supset \boxdot(\varphi \preccurlyeq \psi)$
(A) Absoluteness	for all $w, v \in W$, $f(\varphi, v) = f(\varphi, v)$	$(\varphi \preccurlyeq \psi) \supset \boxdot(\varphi \preccurlyeq \psi)$ + $(\varphi \prec \psi) \supset \boxdot(\varphi \preccurlyeq \psi)$
(UT) Universality	**Uniformity + Total Reflexivity**	**Uniformity + Total Reflexivity**
(WA) Weak Triviality	$f(\varphi, w) = [\varphi]$	**Weak Centering + Absoluteness**
(CA) Triviality	if $w \in [\varphi]$, then $f(\varphi, w) = \{w\}$, otherwise $f(\varphi, w) = \emptyset$	**Centering + Absoluteness**

Table 2: The table schematically summarizes the structural conditions over functional Lewisian models and the corresponding characteristic axioms. The axioms schemata for functional Lewisian models are meant to range over all formulas in $\mathcal{L}_{\boxright}$, and the conditions over all worlds w and all formulas in $\mathcal{L}_{\boxright}$.

We introduce the following notation to refer to classes of functional Lewisian models and define logical consequence with respect to a certain class.

Notation 0.2.

- $\mathbf{V}^f$ *denotes the class of all functional Lewisian models.*
- $\mathfrak{C}$ *indicates a condition or a family of conditions (possibly empty) among those in Table 2, i.e.* $\{\mathbf{N}, \mathbf{T}, \mathbf{W}, \mathbf{C}, \mathbf{S}, \mathbf{A}, \mathbf{U}, \}$.
- $\mathbf{V}^f\mathfrak{C}$ *denotes the class of functional Lewisian models satisfying condition(s)* $\mathfrak{C}$
- *Logical consequence over a class of functional Lewisian models is defined as follows: for* $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}_{\Box\!\rightarrow}}$,

$$\Gamma \models_{\mathbf{V}^f\mathfrak{C}} \varphi \quad \Leftrightarrow \quad \begin{array}{l}\text{for all functional Lewisian models } \mathcal{M} \text{ satisfying conditions } \mathfrak{C},\\ \text{for all } w \text{ in } \mathcal{M}, \text{ if } \mathcal{M}, w \vDash \gamma \text{ for all } \gamma \in \Gamma, \text{ then } \mathcal{M}, w \vDash \varphi\end{array}$$

It is straightforward to check that each axiom characterizes the corresponding condition:

Lemma 0.3. *Let* $\mathcal{M} = \langle W, \mathcal{S}, R, f, [] \rangle$ *be a spherical Lewisian model. We have that*

$$\text{for all } \varphi \in For_{\mathcal{L}_{\Box\!\rightarrow}^{\leq}}, \mathcal{M} \vDash [\text{axiom } \mathfrak{C}] \leftrightarrow \varphi \Leftrightarrow \mathcal{M} \text{ satisfies condition } \mathfrak{C}$$

Functional Lewisian models correspond to α-models in (Lewis 1971). Any spherical Lewisian models (satisfying the limit assumption) is equivalent to a functional Lewisian model. More precisely, as the next result shows, every spherical Lewisian model encodes a functional Lewisian model that validates exactly the same formulas of the original model.

Lemma 0.4. *Any spherical Lewisian model* $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$ *induces a functional Lewisian model* $\mathcal{M}^f = \langle W^f, f, \vDash^f \rangle$ *where*

- $W^f = W$;
- $f : For_{\mathcal{L}_{\Box\!\rightarrow}} \times W \rightarrow \wp(W)$ *is defined as follows: for all* $w \in W$, *for all* $\varphi \in For_{\mathcal{L}_{\Box\!\rightarrow}}$, $f(\varphi, w) = [\varphi] \cap min^{\varphi}_{\subseteq}(\mathcal{S}(w))$;
- $\vDash^f$: $Var \rightarrow \wp(W)$ *is defined as follows: for all* $p \in Var$ *and all* $w \in W$, $w \vDash^f p \Leftrightarrow w \vDash p$ *and* $\vDash^f$ *is extended to compound formulas as in Definition 0.3. We set* $[\varphi] = \{v \in W \mid w \vDash^f\}$.

Moreover, for all $\varphi \in \mathcal{L}_{\Box\!\!\rightarrow}$, it holds that $[\varphi]^f = [\varphi]$, namely that for all $w \in W$, $w \vDash \varphi \Leftrightarrow w \vDash^f \varphi$

Proof. The proof is rather long and tedious and not particularly informative, hence it is included in the appendix (see A.1)[7] □

Lewis (1971) showed that it is also possible to go from functional Lewisian models to spherical Lewisian models by preserving satisfaction:

Lemma 0.5 (Lewis 1971, pp. 78–80)**.** *Any functional Lewisian model $\mathcal{M} = \langle W, f, \vDash\rangle$ induces a spherical Lewisian model $\mathcal{M}^f = \langle W^f, \mathcal{S}^f, \vDash^f\rangle$ such that for all $\varphi \in \mathcal{L}_{\Box\!\!\rightarrow}$, it holds that $[\varphi]^f = [\varphi]$, i.e. for all $w \in W$, $w \vDash \varphi \Leftrightarrow w \vDash^f \varphi$.*

0.4.2 Variably Strict Conditional Logics

Lewis (1971, 1973b) introduced a family of logics, called *variably strict conditional logics*, associated with his models. Those logics can be axiomatized as follows, drawing from (Lewis 1973b):

Definition 0.4. **V** *is the logic induced by the following system of axiom schemata and rules in the language $\mathcal{L}_{\Box\!\!\rightarrow}$:*

- ***Axioms***

(TAUT) *all classical tautologies*

(ID) $\varphi \Box\!\!\rightarrow \varphi$

(CE) $((\varphi \Box\!\!\rightarrow \psi) \wedge (\psi \Box\!\!\rightarrow \varphi)) \supset ((\varphi \Box\!\!\rightarrow \delta) \leftrightarrow (\psi \Box\!\!\rightarrow \delta))$

(NE) $((\varphi \vee \psi) \Box\!\!\rightarrow \varphi) \vee ((\varphi \vee \psi) \Box\!\!\rightarrow \psi) \vee (((\varphi \vee \psi) \Box\!\!\rightarrow \delta)) \leftrightarrow ((\varphi \Box\!\!\rightarrow \delta) \wedge (\psi \Box\!\!\rightarrow \delta)))$

(MIGHT) $(\varphi \Diamond\!\!\rightarrow \psi) \leftrightarrow \neg(\varphi \Box\!\!\rightarrow \neg\psi)$

(BOX) $(\boxdot\varphi) \leftrightarrow (\neg\varphi \Box\!\!\rightarrow \varphi)$

(DIAM) $(\diamonddot\varphi) \leftrightarrow (\neg \boxdot \neg\varphi)$

(PLAUS) $(\varphi \preccurlyeq \psi) \leftrightarrow (((\varphi \vee \psi) \Diamond\!\!\rightarrow (\varphi \vee \psi)) \supset ((\varphi \vee \psi) \Diamond\!\!\rightarrow \varphi))$

- ***Rules***

(MP) *Modus Ponens*

(RTAUT1) $\vdash \varphi$ *when* φ *is a classical tautology*

(RTAUT2) $\varphi_1, \varphi_2, \ldots, \varphi_n \vdash \varphi$ *when* $(\varphi_1 \wedge \cdots \wedge \varphi_n) \supset \varphi$ *is a classical tautology*

[7] It might seem that the proof of this result is superfluous, in the sense that this lemma must intuitively follow from the results in Lewis (1971, 1973b). However, this lemma has never been proved by Lewis. Indeed, an analogous of this lemma is shown in Lewis (1973b, p. 55), however the models employed by Lewis (1973b) are not those we introduced here: our models corresponds to those introduced by Lewis (1971).

(DWC) *for a (possibly empty) set of formulas* $\{\varphi_1, \dots, \varphi_n\}$ $\vdash (\varphi_1 \wedge \cdots \wedge \varphi_n) \supset \varphi$ *then* $\vdash ((\psi \,\Box\!\!\rightarrow \varphi_1) \wedge \cdots \wedge (\psi \,\Box\!\!\rightarrow \varphi_n)) \supset (\psi \,\Box\!\!\rightarrow \varphi)$

(SUB) *substitution of interderivable formulas: if* $\varphi \dashv\vdash_{\mathbf{V}\mathfrak{C}} \psi$, *then* $\delta \dashv\vdash_{\mathbf{V}\mathfrak{C}} \delta[\varphi/\psi]$

- *For an axiom or a family of axioms (possibly empty)* $\mathfrak{C}$ *among those in Table 2 (or Table 1),* $\mathbf{V}\mathfrak{C}$ *is the logic induced by the system obtained by extending* **V** *with the axiom(s) in* $\mathfrak{C}$*. For instance,* **VC** *is the logic induced by adding the axiom* **C** *to* **V**.
- *For* $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}_{\Box\rightarrow}}$*, a derivation from* Γ *to* φ *in* $\mathbf{V}\mathfrak{C}$ *is a finite sequence of formulas that ends with* φ *such that each formula in the sequence instantiates an axiom in* $\mathbf{V}\mathfrak{C}$*, or belongs to* Γ*, or it is obtained by applications of the rules in* $\mathbf{V}\mathfrak{C}$*.* $\Gamma \vdash_{\mathbf{V}\mathfrak{C}} \varphi$ *means that* φ *is derivable from* Γ *in* $\mathbf{V}\mathfrak{C}$*.*
- *For* $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}_{\Box\rightarrow}}$*,* $\Gamma \vdash_{\mathbf{V}\mathfrak{C}} \varphi$ *means that there is a derivation from formulas in* Γ *to* φ*.*
- *From the definition of derivation, it is straightforward to see that for* $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}_{\Box\rightarrow}}$*,*

$$\Gamma \vdash_{\mathbf{V}\mathfrak{C}} \varphi \Leftrightarrow \Delta \vdash_{\mathbf{V}\mathfrak{C}} \varphi \text{ for some finite } \Delta \subseteq \Gamma$$

The following soundness and completeness results hold, by easily adapting the soundness and completeness proofs in (Lewis 1971, 1973b), in addition to Lemmas 0.5 and 0.4:

Theorem 0.1. *Let* $\mathfrak{C}$ *be an axiom/condition or a family of axioms/conditions (possibly empty) among those in Table 1, (i.e.* $\{\mathbf{N}, \mathbf{T}, \mathbf{W}, \mathbf{C}, \mathbf{S}, \mathbf{A}, \mathbf{U},\}$*), then, for all* $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}_{\Box\rightarrow}}$*, the following holds:*

$$\Gamma \vdash_{\mathbf{V}\mathfrak{C}} \varphi \Leftrightarrow \Gamma \models_{\mathbf{V}\mathfrak{C}} \varphi \Leftrightarrow \Gamma \models_{\mathbf{V}^{\mathfrak{f}}\mathfrak{C}} \varphi$$

Notably, among the variably strict conditional logics, **VC** (=**C1** in Lewis 1971) corresponds to the "*correct logic of counterfactual conditionals, as we ordinarily understand them*", according to Lewis (1971). Thus, with the expression "Lewis counterfactuals", we refer to those conditionals obeying the rules and axioms in the logic **VC**. The logic **VCS** (=**C2** in Lewis 1971) corresponds to Stalnaker's logic of conditionals; thus with the expression "Stalnaker conditionals", we refer to those conditionals obeying the rules and axioms in the logic **VCS**. By looking at Table 2 we can see that the two accounts differ in that Lewis **VC** models allow for the selection function to contain more than one world, whereas Stalnakerian **VCS** models constraint the selection function to contain *at most* on world. This reflects on the logics: $\models_{\mathbf{VCS}} (\varphi \mathrel{\Box\!\!\rightarrow} \psi) \vee (\varphi \mathrel{\Box\!\!\rightarrow} \neg\psi)$ whereas $\not\models_{\mathbf{VC}} (\varphi \mathrel{\Box\!\!\rightarrow} \psi) \vee (\varphi \mathrel{\Box\!\!\rightarrow} \neg\psi)$.

We have now all the background ingredients to move to the next chapters focusing on our main original contributions.

Chapter 1

The Algebra of Counterfactuals

In this chapter, we embark on an algebraic treatment of Lewis' variably strict conditional logics, also known as **V**-logics, which we introduced in the previous chapter. Specifically,we will address questions (A1), (A1a), and (A1b) mentioned in the introduction. While Nute (1975) and Weiss (2019) have previously developed an algebraic semantics for variably strict conditional logic, their approach lacks the spirit of abstract algebraic logic initiated by Blok and Pigozzi (2014). Their work only establishes a weak completeness of variably strict conditional logics with respect to special algebraic structures.

However, a completeness result represents merely an initial step towards achieving a comprehensive and systematic algebraic treatment of a family of logics. In the following, we take a *further* step in the direction of developing such a systematic algebraic approach to variably strict conditional logics. Our goal is to show that a special class of algebraic structures are not only complete with respect to **V**-logics, but also serve as an *adequate* algebraic counterpart of **V**-logics. In other words, we aim to show that **V**-logics are algebrizable, in the sense of Blok and Pigozzi (2014), with respect to the class of algebras we introduce.

By doing so, we strive to contribute to the broader field of abstract algebraic logic and provide a more robust foundation for understanding and analyzing variably strict conditional logics. This will enable us to address questions and challenges related to their algebraic aspects more effectively and pave the way for further advancements in this area of research.

1.1 Global vs Local Consequence

In this section, we undertake a comprehensive logical examination of Lewis' variably strict conditional logics, focusing solely on the conceptual aspects and abstaining from introducing any algebraic notions at this stage. Our objective is to delve deeper into the properties of Lewis' variably strict conditional logic, drawing insightful analogies with modal logic and its model-theoretic properties, as extensively studied by Blackburn, de Rijke, and Venema (2001).

Consider a classical Kripke model for modal logic $\langle W, R, \vDash \rangle$. Formulas in a model can be interpreted *locally* or *globally*. A formula φ is locally true at $\mathcal{M}$ if it is true at some world in W, i.e. for some $w \in W$, $\mathcal{M}, w \vDash \varphi$. On the other hand, φ is globally true at $\mathcal{M}$ if and only if it is true all worlds in W, i.e. if for all $w \in W$, $\mathcal{M}, w \vDash \varphi$. The same distinctions can be drawn inside a spherical (or functional) Lewisian model for the language $\mathcal{L}_{\Box\!\!\rightarrow}$.

Definition 1.1. *Consider a spherical Lewisian model $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$; we say that a formula $\varphi \in For_{\mathcal{L}_{\Box\!\!\rightarrow}}$ is:*

- ***locally true** at $\mathcal{M}$ if it is true at some world w of W. We write $\mathcal{M}, w \vDash \varphi$ to indicate that φ is true at w.*
- ***globally true** at $\mathcal{M}$ if it is true at all worlds in W. We write $\mathcal{M} \vDash \varphi$ to indicate that for all $w \in W$, $\mathcal{M}, w \vDash \varphi$*

Global and local truth naturally induce corresponding notions of logical consequence:

Definition 1.2 (Global vs Local)**.**

- ***Local.** The local logical consequence relation is defined as follows: for all $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}_{\Box\!\!\rightarrow}}$,*

$$\Gamma \models^{l}_{\mathbf{V}\mathfrak{C}} \varphi \quad \Leftrightarrow \quad \begin{array}{l}\textit{for all spherical Lewisian models } \mathcal{M} \textit{ satisfying conditions } \mathfrak{C},\\ \textit{for all } w \textit{ in } \mathcal{M}, \textit{ if } \mathcal{M}, w \vDash \gamma \textit{ for all } \gamma \in \Gamma, \textit{ then } \mathcal{M}, w \vDash \varphi\end{array}$$

- ***Global.** The global logical consequence relation is defined as follows: for all $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}_{\Box\!\!\rightarrow}}$,*

$$\Gamma \models^{g}_{\mathbf{V}\mathfrak{C}} \varphi \quad \Leftrightarrow \quad \begin{array}{l}\textit{for all spherical Lewisian models } \mathcal{M} \textit{ satisfying conditions } \mathfrak{C},\\ \textit{if } \mathcal{M} \vDash \gamma \textit{ for all } \gamma \in \Gamma, \textit{ then } \mathcal{M} \vDash \varphi\end{array}$$

The local consequence relation $\models^{l}_{\mathbf{V}\mathfrak{C}}$ is just a notational variant of the notion of logical consequence $\models_{\mathbf{V}\mathfrak{C}}$ introduced in Section 0.4. We introduced a different notation with the superscript ^{l} to highlight the distinction between global and local consequences.

It is clear that Lewis' variably strict conditional logics coincide with the notion of local logical consequence over spherical Lewisian model. The difference between the two notions can be further appreciated from a proof theoretic perspective.

Definition 1.3 (Global **V**-logics). *The global* $\mathbf{V}\mathfrak{C}$*-logic is the logic induced by the axiom system obtained by adding the following rules to the corresponding* $\mathbf{V}\mathfrak{C}$ *system in Definition 0.4:*

(DWCG) $(\varphi_1 \wedge \cdots \wedge \varphi_n) \to \psi \vdash_{\mathbf{C1}_g} ((\gamma \Box\!\!\rightarrow \varphi_1) \wedge \cdots \wedge (\gamma \Box\!\!\rightarrow \varphi_n)) \to (\gamma \Box\!\!\rightarrow \psi)$

(SUBG) $\varphi \leftrightarrow \psi \vdash^g_{\mathbf{V}\mathfrak{C}} \delta \leftrightarrow \delta[\varphi/\psi]$

We use the symbol $\vdash^g_{\mathbf{V}\mathfrak{C}}$ *to indicate derivation in the global* $\mathbf{V}\mathfrak{C}$*-logic. In particular, for all* $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}_{\Box\!\rightarrow}}$*,* $\Gamma \vdash^g_{\mathbf{V}\mathfrak{C}} \varphi$ *means that there is a derivation in the global* $\mathbf{V}\mathfrak{C}$*-logic from formulas in* Γ *to* φ

Observe that by adding $(DWCG)$ and $(SUBG)$ to our logical systems, the old rules (DWG) and (SUB) become redundant.

Notation 1.1.

- *The local* $\mathbf{V}\mathfrak{C}$*-logic,* $\vdash^l_{\mathbf{V}\mathfrak{C}}$ *is exactly the logic* $\mathbf{V}\mathfrak{C}$ *introduced in Definition 0.4.*
- *We use the symbol* $\vDash^l_{\mathbf{V}\mathfrak{C}}$ *to indicate derivation in the local* $\mathbf{V}\mathfrak{C}$*-logic. Specifically, for all* $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}_{\Box\!\rightarrow}}$*, we set:*

$$\Gamma \vDash^l_{\mathbf{V}\mathfrak{C}} \varphi \Leftrightarrow \Gamma \vdash_{\mathbf{V}\mathfrak{C}} \varphi$$

- *The local* $\mathbf{V}\mathfrak{C}$*-logic* $\vdash^l_{\mathbf{V}\mathfrak{C}}$ *is just a notational variant of* $\mathbf{V}\mathfrak{C}$*-logic from Definition 1.3.*

From a proof-theoretic perspective, we can already observe some connections and and some differences between local and global $\mathbf{V}\mathfrak{C}$-logics.

Remark 1.1. *The following hold:*

1. $(\varphi_1 \wedge \cdots \wedge \varphi_n) \supset \psi \vdash^g_{\mathbf{V}\mathfrak{C}} ((\gamma \Box\!\!\rightarrow \varphi_1) \wedge \cdots \wedge (\gamma \Box\!\!\rightarrow \varphi_n)) \supset (\gamma \Box\!\!\rightarrow \psi)$
2. $(\varphi_1 \wedge \cdots \wedge \varphi_n) \supset \psi \nvdash^l_{\mathbf{V}\mathfrak{C}} ((\gamma \Box\!\!\rightarrow \varphi_1) \wedge \cdots \wedge (\gamma \Box\!\!\rightarrow \varphi_n)) \supset (\gamma \Box\!\!\rightarrow \psi)$
3. *for all* $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}_{\Box\!\rightarrow}}$*,*
$$\Gamma \vdash^l_{\mathbf{V}\mathfrak{C}} \varphi \Rightarrow \Gamma \vdash^g_{\mathbf{V}\mathfrak{C}} \varphi$$
4. *The local* $\mathbf{V}\mathfrak{C}$*-logic and the global* $\mathbf{V}\mathfrak{C}$*-logic share the same theorems: for all* $\varphi \in For_{\mathcal{L}_{\Box\!\rightarrow}}$*,*
$$\vdash^g_{\mathbf{V}\mathfrak{C}} \varphi \Leftrightarrow \vdash^l_{\mathbf{V}\mathfrak{C}} \varphi$$

Proof.

1. Straightforward by rule $(DWCG)$ in the system of global $\mathbf{V}\mathfrak{C}$;

2. Semantically, we have that

$$(\varphi_1 \wedge \cdots \wedge \varphi_n) \to \psi \not\models^{l}_{\mathbf{V}\mathfrak{C}} ((\gamma \mathrel{\Box\!\!\rightarrow} \varphi_1) \wedge \cdots \wedge (\gamma \mathrel{\Box\!\!\rightarrow} \varphi_n)) \supset (\gamma \mathrel{\Box\!\!\rightarrow} \psi)$$

   Indeed, it is very easy to find a counter-model for this inference. Furthermore, since by Theorem 0.1 the local $\mathbf{V}\mathfrak{C}$-logic coincides with $\models_{\mathbf{V}\mathfrak{C}}$, we have that

$$(\varphi_1 \wedge \cdots \wedge \varphi_n) \to \psi \nvdash^{l}_{\mathbf{V}\mathfrak{C}} ((\gamma \mathrel{\Box\!\!\rightarrow} \varphi_1) \wedge \cdots \wedge (\gamma \mathrel{\Box\!\!\rightarrow} \varphi_n)) \supset (\gamma \mathrel{\Box\!\!\rightarrow} \psi)$$

3. Straightforward by observing that all the axioms and rules of the local $\mathbf{V}\mathfrak{C}$-logic are also axioms and rules of the global $\mathbf{V}\mathfrak{C}$-logic.

4. $(\Leftarrow)$ Straightforward by point 3 above.

   $(\Leftarrow)$ We provide a proof-sketch. If $\vdash^{g}_{\mathbf{V}\mathfrak{C}} \varphi$, then either φ is an axiom or it is derived by assuming only theorems and applying rules of the system $\mathbf{V}\mathfrak{C}$. If φ is an axiom of global $\mathbf{V}\mathfrak{C}$, then also $\vdash^{l}_{\mathbf{V}\mathfrak{C}}$ since the global and local systems share the same axiom. In the other case, we have that $\Gamma \vdash^{g}_{\mathbf{V}\mathfrak{C}} \varphi$ and Γ contain only axioms or theorems of the global $\mathbf{V}\mathfrak{C}$. Observe that every proof of φ in global $\mathbf{V}\mathfrak{C}$ from such a Γ is also a proof in local $\mathbf{V}\mathfrak{C}$. Indeed, the only difference might be in the fact that the derivation $\Gamma \vdash^{g}_{\mathbf{V}\mathfrak{C}} \varphi$ employs $(DWCG)$. However, since all formulas in Γ are axioms of local $\mathbf{V}\mathfrak{C}$ too, every application of $(DWCG)$ in $\Gamma \vdash^{g}_{\mathbf{V}\mathfrak{C}} \varphi$ is basically an application of (DWC). Hence the same derivation is also a derivation in local $\mathbf{V}\mathfrak{C}$, that is: $\Gamma \vdash^{l}_{\mathbf{V}\mathfrak{C}} \varphi$

□

In the following, we analyze the differences and the connections between the global an the local systems from a semantic perspective.

1.1.1 Invariance Results

To the best of our knowledge, the distinction between the local and global consequence relation in variably strict conditional logic has not been previously addressed in the relevant literature. However, an analogous differentiation exists in the case of modal logic. Blackburn, de Rijke, and Venema (2001) differentiate a local modal logic, which corresponds to the usual normal modal logic, from a *global* modal logic which, that emerges when the restriction of the necessitation rule is lifted. That is, in the global modal logics we have $\varphi \vdash \Box\varphi$, while in the local counterpart we have a restricted version of the rule: if $\vdash \varphi$, then $\vdash \Box\varphi$. Notably, these two logics also possess semantic counterparts: the global modal logic coincides with the global truth-preserving consequence over Kripke models, while the local consequence corresponds the standard local truth-preserving consequence over Kripke models. The connections between the global and local consequence in modal logic have been explored by Kracht (1999) and Wen (2021). Paralleling this distinction in modal logic, we will investigate the relationship between the local and global consequence within the context of spherical Lewisian models. To initiate this examination, Definition 1.2 serves as a fundamental starting point, facilitating the following observation:

Remark 1.2. *For all* $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}_{\Box\rightarrow}}$*,*

1. $\models^g_{\mathbf{V}\mathfrak{C}} \varphi \Leftrightarrow \models^l_{\mathbf{V}\mathfrak{C}} \varphi$
2. $\Gamma \models^l_{\mathbf{V}\mathfrak{C}} \varphi \Rightarrow \Gamma \models^g_{\mathbf{V}\mathfrak{C}} \varphi$

Namely local and global consequences share the same validities; additionally, local consequence is sound with respect to global consequence:

There is a significant and intricate connection between local and global consequence relations. Specifically, the global consequence relation can be effectively characterized using the local one. However, before presenting this characterization result, it is necessary to introduce a valuable tool that resembles the concept of a *generated sub-model*, similar to what Blackburn, de Rijke, and Venema (2001) employed in the context of Kripke models. Drawing a parallel with Kripke semantics, we will demonstrate how to manipulate a sphere model to create a new model that preserves the satisfaction of formulas. This manipulation will play a crucial role in proving a key *invariance result* for Lewis semantics of counterfactuals.

Definition 1.4. *Let* $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$ *and* $\mathcal{M}' = \langle W', \mathcal{S}', \vDash' \rangle$ *be two spherical Lewisian models. We say that* $\mathcal{M}'$ *is a sub-model of* $\mathcal{M}$ *if the following hold:*

- $W' \subseteq W$
- $\mathcal{S}'$ *is the restriction of* $\mathcal{S}$ *to* W'*, i.e.* $\mathcal{S}' : W' \to \wp(\wp(W'))$ *and for all* $w' \in W'$*,* $\mathcal{S}'(w') = \mathcal{S}(w) \cap \wp(\wp(W'))$.
- $\vDash'$ *is the restriction of* $\vDash$ *to* W'*, i.e. for all* $p \in Var$*, for all* $w' \in W'$*,* $w' \vDash' p \Leftrightarrow w' \vDash p$*.* $\vDash'$ *is extended to all formulas in* $For_{\mathcal{L}_{\Box\rightarrow}}$ *according to the clauses in Definition 0.2.*

A special class of sub-models will prove extremely useful, namely *generated sub-models*.

Definition 1.5. *Let $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$ and $\mathcal{M}' = \langle W', \mathcal{S}', \vDash' \rangle$ be two spherical Lewisian models. We say that $\mathcal{M}'$ is a generated sub-model of $\mathcal{M}$ if and only if $\mathcal{M}'$ is a sub-model of $\mathcal{M}$ and moreover the following condition holds: for all $w', v' \in W$,*

$$\textit{if } w' \in W' \textit{ and } v' \in \bigcup \mathcal{S}(w'), \textit{ then } v' \in W'$$

Namely, for all $w' \in W'$, $\mathcal{S}'(w') = \mathcal{S}(w')$.

Moreover, if the original model $\mathcal{M}$ satisfies a constraint $\mathfrak{C}$ among those in Table 1, its generated sub-model $\mathcal{M}'$ satisfies constraint $\mathfrak{C}$ too.

In other words, a generated sub-model consist in selecting some worlds $w_1, w_2, \ldots$ from the original model, essentially forming a a sub-domain of the original model. However, we must adhere to the proviso that for all worlds w_i we selected, their systems of spheres must be carried over into the new generated sub-model. This particular type of generated sub-model will play a key role in our analysis:

Definition 1.6. *Let $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$ and consider $X \subseteq W$; the sub-model generated by X is the smallest (w.r.t. set inclusion, i.e. the size of their underlying domains) sub-model of $\mathcal{M}$ whose domain contains X. Moreover, a centered or point generated sub-model of $\mathcal{M}$ is a sub-model of $\mathcal{M}$ generated by a singleton set.*

The above definition may seem odd since we have not proved that such *smallest* sub-model exists. In what follows, we fill this gap and provide a more informative characterization of generated sub-models.

Definition 1.7. *Consider a spherical Lewisian model $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$ and a subset $X \subseteq W$. Now, define a binary relation $R \subseteq W \times W$ as follows: for all $w, v \in W$,*

$$wRv \Leftrightarrow v \in \bigcup \mathcal{S}(w)$$

Namely, wRv if and only if v appears in the system of spheres associated to w. That is to say that every worlds accesses to all the worlds in its systems of spheres. Now, for all $n \in \mathbb{N}$, define a relation $R^n \subseteq W \times W$ by induction in the following way:

- $wR^0v \Leftrightarrow w = v$
- $wR^{n+1}v \Leftrightarrow$ *there is a $u \in W$ such that wR^nu and uRv.*

Now, consider the subset $M \subseteq W$ defined as follows:

$$M = \{v \in W \mid \textit{ there is a } n \in \mathbb{N} \textit{ and there is a } w \in X \textit{ such that } wR^nv\}$$

Namely M is the set made of all the elements of X and all the worlds in W that are reachable from a member of X by a finite number of steps via R.
Now, call $\mathcal{M}' = \langle M, \mathcal{S}', \vDash' \rangle$ the spherical Lewisian model such that:

- *$\mathcal{S}'$ is the restriction of $\mathcal{S}$ to M*
- *$\vDash'$ is the restriction of $\vDash$ to M.*

We refer to R in the above definition as the accessibility relation induced by $\mathcal{S}$. Intuitively, the relation $wR^n v$ above indicates the number of steps needed to reach the world v starting from w. For instance, for every $w \in IW$, we have $wR^0 w$, that is, "no step" is needed to reach w from itself; $wR^4 v$ means that there are worlds m, k, l such that wRm, mRk, kRl and lRv. We are now ready to prove the following;

Lemma 1.1. *Consider a spherical Lewisian model $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$ and the sub-model $\mathcal{M}'$ as in Definition 1.7. Then $\mathcal{M}'$ is the smallest generated sub-model of $\mathcal{M}$ whose domain contains X, that is $\mathcal{M}'$ is the smallest sub-model of $\mathcal{M}$, which is generated by X .*

Proof. First, observe that $\mathcal{M}'$ is a well defined sub-model of $\mathcal{M}'$. Moreover, it is also a generated sub-model of $\mathcal{M}$. In order to see that, observe that the following condition holds in $\mathcal{M}'$:

$$\text{if } w \in M \text{ and } wRv, \text{ then } v \in M$$

Indeed, assume $w \in M$ and wRv for some $v \in W$, this means that $wR^1 v$ by definition of R^n, and so $v \in M$. Observe that the above condition, by definition of R, can be rewritten as:

$$\text{if } w \in M \text{ and } v \in \bigcup \mathcal{S}(w), \text{ then } v \in M$$

Hence, $\mathcal{M}'$ is a generated sub-model of $\mathcal{M}$ by Definition 1.5. Now, consider any generated sub-model, $\mathcal{M}^* = \langle W^*, \mathcal{S}^*, \vDash^* \rangle$ whose domain contains X, i.e. $X \subseteq W^*$. We are going to show that $M \subseteq W^*$. In order to do that, we will prove the following claim, namely that for all $w \in W$, for all $n \in \mathbb{N}$ if there is a $x \in X$ such that $xR^n w$, then $w \in W^*$. We prove it by induction on n:

- if $n = 0$, then $x = x$ and so $x \in X$, and since $X \subseteq W^*$, we also have that $x \in W^*$;
- if $n = m + 1$, then there is a $x \in X$ such that $xR^{m+1} w$, that is to say that there is a $l \in W$ such that $xR^m l$ and lRw, By Induction hypothesis, we have that $l \in W^*$. Since $\mathcal{M}^*$ is a generated sub-model of $\mathcal{M}$, we have that since $l \in W^*$ and lRw, i.e. $w \in \bigcup \mathcal{S}(l)$, it is also the case that $w \in W^*$.

Now, consider a $m \in M$; by definition of M, there is a $x \in X$ and $n \in \mathbb{N}$ such that $xR^n m$ and so, by what we proved above, it holds that $m \in W^*$, and so $M \subseteq W^*$. Namely, $\mathcal{M}'$ is the smallest sub-model of $\mathcal{M}$ generated by X. □

Generated sub-models preserves truth of formulas, as the following lemmas show:

Lemma 1.2. *Let $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$ be a spherical Lewisin model, and let $\mathcal{M}' = \langle W', \mathcal{S}', \vDash' \rangle$ be a generated sub-model of $\mathcal{M}$. The following holds: for all $w \in W'$, for all formulas φ,*

$$\mathcal{M}, w \vDash \varphi \Leftrightarrow \mathcal{M}', w \vDash' \varphi$$

Proof. By induction on φ: we will prove two cases for exemplification.

1. $\varphi = p$. Clearly by definition of generated sub-model since $\vDash'$ is the restriction of $\vDash$ to W';

2. $\varphi = \psi \mathrel{\Box\!\!\rightarrow} \delta$. Assume $\mathcal{M}, w \vDash \psi \mathrel{\Box\!\!\rightarrow} \delta$; observe that, by definition of sub-model, for all $w \in W'$, $\mathcal{S}(w) = \mathcal{S}'(w)$. Hence, we have two cases to consider. Let us use the following notation: for all $\varphi \in For_{\mathcal{L}_{\Box\!\rightarrow}}$, $[\varphi]' = \{w' \in W' \mid w' \vDash' \varphi\}$.

   - $\bigcup \mathcal{S}(w) \cap [\psi] = \emptyset$; so, since $\vDash'$ and $\mathcal{S}'$ are the restrictions of $\mathcal{S}$ and $\vDash$ to W', we will also have that $\bigcup \mathcal{S}'(w) \cap [\psi]' = \emptyset$, and so $\mathcal{M}', w \vDash' \psi \mathrel{\Box\!\!\rightarrow} \delta$.
   - there is a $S \in \mathcal{S}(w)$ such that $[\psi] \cap S \neq \emptyset$ and $([\psi] \cap S) \subseteq [\delta]$. Again, since $\vDash'$ and $\mathcal{S}'$ are the restrictions of $\mathcal{S}$ and $\vDash$ to W', we also have that there is a $S' \in \mathcal{S}'(w)$ such that $[\psi]' \cap S' \neq \emptyset$ and $([\psi]' \cap S') \subseteq [\delta]'$. And so $\mathcal{M}', w \vDash' \psi \mathrel{\Box\!\!\rightarrow} \delta$

   We reason analogously for the other direction.

□

We are approaching our main results, that is the characterization of global consequence via local consequence. At this point, a new notation can prove useful. Recall the definition of the $\boxdot$ operator from Definition 0.1: $\boxdot\varphi$: $\neg\varphi \mathrel{\Box\!\!\rightarrow} \varphi$:

Notation 1.2.

- *for all $n \in \mathbb{N}$, $\boxdot^n\varphi$ is inductively defined as follows:*
  - $\boxdot^0\varphi := \varphi$

– *for* $n > 0$, $\boxdot^n \varphi := \boxdot \boxdot^{n-1} \varphi$

namely $\boxdot^n \varphi$ *is the formula consisting in* φ *preceded by a n number of* $\boxdot$ *operators.*

- *for instance,* $\boxdot^3 = \boxdot \boxdot \boxdot \varphi$

Now, observe the following:

Lemma 1.3. *For any spherical Lewisian model* $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$*, the following holds: let* R *be defined as in Lemma 1.1, then for all* $w \in W$*, for all* $\varphi \in For_{\mathcal{L}^{\square}_{>}}$*,*

$$w \vDash \boxdot\varphi \Leftrightarrow w \vDash \bigcup \mathcal{S}(w) \Leftrightarrow \textit{for all } v : wRv,\, v \vDash \varphi$$

Proof. Observe that $w \vDash \neg\varphi \;\square\!\!\rightarrow\; \varphi$ if and only if $\bigcup \mathcal{S}(w) \cap [\neg\varphi] = \emptyset$, and this means that $\bigcup \mathcal{S}(w) \cap [\varphi] = \bigcup \mathcal{S}(w)$, and so $\bigcup \mathcal{S}(w) \subseteq [\varphi]$. Given this observation, the above lemma follows straightforwardly by definition of $\boxdot$ and R. □

Finally, the following theorem establishes a deep connection between local and global consequences:

Theorem 1.1. *For al* $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}_{\square\!\rightarrow}}$*,*

$$\Gamma \vDash^{g}_{\mathbf{V}\mathfrak{C}} \varphi \Leftrightarrow \{\boxdot^n \gamma \mid n \in \mathbb{N} \textit{ and } \gamma \in \Gamma\} \vDash^{l}_{\mathbf{V}\mathfrak{C}} \varphi$$

Proof.

(⇒) By contrapositio, assume $\Gamma \not\vDash^{g}_{\mathbf{V}\mathfrak{C}} \varphi$. Then there is a spherical Lewisian model $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$ satisfying condition $\mathfrak{C}$ such that for all $w \in W$, $w \vDash \gamma$, for all $\gamma \in \Gamma$, and for some $v \in W$, $v \nvDash \varphi$. Now, it is easy to see that for all $n \in \mathbb{N}$ and all $w \in W$, $w \vDash \boxdot^n \gamma$ for all $\gamma \in \Gamma$. We prove it by induction on n: consider $w \in W$

  – if $n = 0$, then $w \vDash \gamma$ for all $\gamma \in \Gamma$ by assumption, since all the worlds in W verify all the formulas in Γ;

  – if $n = m + 1$, if $\bigcup \mathcal{S}(w) = \emptyset$, then clearly $w \vDash \boxdot^{m+1}\gamma$ for all $\gamma \in \Gamma$; if $\bigcup \mathcal{S}(w) \neq \emptyset$, then consider any $y \in \bigcup \mathcal{S}(w)$; by induction hypothesis $y \vDash \boxdot^m \gamma$ for all $\gamma \in \Gamma$. This means that $w \vDash \boxdot \boxdot^m \gamma$ for all $\gamma \in \Gamma$, that is $w \vDash \boxdot^{m+1}\gamma$ for all $\gamma \in \Gamma$.

So, in particular, for all $n \in \mathbb{N}$, including $n = 0$ and all $\gamma \in \Gamma$, we have that $v \vDash \boxdot^n \gamma$, and so $\{\boxdot^n \gamma \mid n \in \mathbb{N} \text{ and } \gamma \in \Gamma\} \not\vDash^{l}_{\mathbf{V}\mathfrak{C}} \varphi$ since $v \vDash \boxdot^n \gamma$ for all $n \in \mathbb{N}$ and all $\gamma \in \Gamma$, but $v \nvDash \varphi$ by assumption.

($\Rightarrow$) By contrapositio, assume $\{\boxdot^n\gamma \mid n \in \mathbb{N} \text{ and } \gamma \in \Gamma\} \nvDash^l_{\mathbf{V}\mathfrak{C}} \varphi$; so there is a spherical Lewisian model $\mathcal{M} = \langle W, \mathcal{S}, \vDash\rangle$ satisfying condition $\mathfrak{C}$ such that for some $w \in W$, $w \vDash \boxdot^n\gamma$ for all $n \in \mathbb{N}$ and all $\gamma \in \Gamma$ and $w \nvDash \varphi$. Now, consider the sub-model $\mathcal{M}' = \langle W', \mathcal{S}', \vDash'\rangle$ generated by $\{w\}$. By Definition 1.5, we have that $\mathcal{M}'$ is also a spherical Lewisian model satisfying condition $\mathfrak{C}$. By Lemma 1.2, we have that for all $w' \in W'$ $\mathcal{M}, w' \vDash \varphi$ if and only if $\mathcal{M}', w' \vDash' \varphi$. Hence, in particular, $\mathcal{M}', w \vDash' \boxdot^n\gamma$ for all $n \in \mathbb{N}$ and all $\gamma \in \Gamma$ and $\mathcal{M}', w \nvDash' \varphi$. Now, we will prove that all the worlds $w' \in W'$ are such that $w' \vDash \boxdot^n\gamma$ for all $\gamma \in \Gamma$ and $n \in \mathbb{N}$. In order to do that, we will rely on Lemma 1.1 and Lemma 1.3 and prove that for all $v \in W$ such that $wR^m v$ for some $m \in \mathbb{N}$, it is the case that $v \vDash \boxdot^n\gamma$ for all $\gamma \in \Gamma$ and all $n \in \mathbb{N}$. Namely, all the R^m accessible worlds from w forces the $\boxdot^n\gamma$'s. We prove by induction on m that for any $v \in W$ such that $wR^m v$, $v \vDash \boxdot^n\gamma$ for all $\gamma \in \Gamma$ and all $n \in \mathbb{N}$

- if $m = 0$, then $v = w$ and so, by assumption, $w \vDash \boxdot^n\gamma$ for all $\gamma \in \Gamma$ and $n \in \mathbb{N}$
- if $m = r + 1$, then, by Lemma 1.1, there is a $k \in W'$ such that $wR^r k$ and kRv. By induction hypothesis, we have that $k \vDash \boxdot^n\gamma$ for all $\gamma \in \Gamma$ and $n \in \mathbb{N}$. Now, consider any $n \in \mathbb{N}$. By induction hypothesis, it is the case that $k \vDash \boxdot^{n+1}\gamma$ for all $\gamma \in \Gamma$, namely, $k \vDash \boxdot\boxdot^n\gamma$ for all $\gamma \in \Gamma$. Since kRv, by Lemma 1.3, then $v \vDash \boxdot^n\gamma$, and this holds for all n.

Now, consider any $x \in W'$, it is the case that for some $m \in \mathbb{N}$, $wR^m x$, by definition of $\mathcal{M}'$ and Lemma 1.1. Thus, by the results we proved above and Lemma 1.2, we have that for all $x \in W'$, $x \vDash' \boxdot^n\gamma$ for all $n \in \mathbb{N}$ and all $\gamma \in \Gamma$. But w is such that $w \nvDash' \varphi$. This implies that $\Gamma \nvDash^g_{\mathbf{V}\mathfrak{C}} \varphi$.

□

In the next section, we will see how to employ the above characterization result to show some properties of the global $\mathbf{V}\mathfrak{C}$-logic.

1.1.2 Completeness and Deduction Theorem

In this section, we are going to prove completeness of global $\mathbf{V}\mathfrak{C}$-logics with respect to global consequence over the class of $\mathbf{V}\mathfrak{C}$ spherical Lewisian models. To begin, some properties of $\mathbf{V}\mathfrak{C}$ logical systems will be useful.

Lemma 1.4. *For any global $\mathbf{V}\mathfrak{C}$-logic, the following rules are derivable:*

(NEC) $\varphi \vdash^g_{\mathbf{V}\mathfrak{C}} \boxdot\varphi$

(*NECn*) $\varphi \vdash^{g}_{\mathbf{V}\mathfrak{C}} \boxdot^{n}\varphi$ *for all* $n \in \mathbb{N}$

- *Moreover, the following holds: for all* $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}_{\Box\mapsto}}$,

$$\Gamma \vdash^{g}_{\mathbf{V}\mathfrak{C}} \varphi \Leftrightarrow \{\boxdot^{n}\gamma \mid \gamma \in \Gamma \textit{ and } n \in \mathbb{N}\} \vdash^{g}_{\mathbf{V}\mathfrak{C}} \varphi$$

Proof.

(NEC) Straightforward by definition of $\boxdot$ and the rule ($DWCG$).

($NECn$) By iterated application of the rule (NEC).

- ($\Rightarrow$) Straightforward by weakening, since $\Gamma \subseteq \{\boxdot^{n}\gamma \mid \gamma \in \Gamma$ and $n \in \mathbb{N}\}$.

  ($\Leftarrow$) Assume $\{\boxdot^{n}\gamma \mid \gamma \in \Gamma$ and $n \in \mathbb{N}\} \vdash^{g}_{\mathbf{V}\mathfrak{C}} \varphi$ and consider Γ. By ($NECn$), we have that $\Gamma \vdash^{g}_{\mathbf{V}\mathfrak{C}} \boxdot^{n}\gamma$ for all $\gamma \in \Gamma$ and $n \in \mathbb{N}$. Moreover, by assumption, from all the $\boxdot^{n}\gamma$'s we can derive φ. Hence, by transitivity of $\vdash^{g}_{\mathbf{V}\mathfrak{C}}$, $\Gamma \vdash^{g}_{\mathbf{V}\mathfrak{C}} \varphi$

□

Now, we have all the ingredients to prove our soundness and completeness results:

Theorem 1.2. *For all* $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}_{\Box\mapsto}}$,

$$\Gamma \vdash^{g}_{\mathbf{V}\mathfrak{C}} \varphi \Leftrightarrow \Gamma \models^{g}_{\mathbf{V}\mathfrak{C}} \varphi$$

Proof.

($\Rightarrow$) It easy to show that all the axioms of the global $\mathbf{V}\mathfrak{C}$-logic are valid with respect to $\models^{g}_{\mathbf{V}\mathfrak{C}}$, and that rules in the global $\mathbf{V}\mathfrak{C}$-logic preserve global truth of formula.s

($\Leftarrow$) By contraposition, assume $\Gamma \nvdash^{g}_{\mathbf{V}\mathfrak{C}} \varphi$. By Lemma 1.4, we have that $\{\boxdot^{n} \mid \gamma \in \Gamma$ and $n \in \mathbf{N}\} \nvdash^{g}_{\mathbf{V}\mathfrak{C}} \varphi$. By Remark 1.1 and Theorem 0.1, we have that $\{\boxdot^{n}\gamma \mid \gamma \in \Gamma$ and $n \in \mathbf{N}\} \not\models^{l}_{\mathbf{V}\mathfrak{C}} \varphi$. Now, by Theorem 1.1, we have that $\Gamma \not\models^{g}_{\mathbf{V}\mathfrak{C}} \varphi$.

□

However, global $\mathbf{V}\mathfrak{C}$ consequence doesn't have a *standard* deduction theorem with respect to material implication. In particular:

Remark 1.3.

1. $p \models^{g}_{\mathbf{V}\mathfrak{C}} \boxdot p$ *, but* $\not\models^{g}_{\mathbf{V}\mathfrak{C}} p \supset \boxdot p$

2. *whereas, for all* $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}_{\Box\rightarrow}}$,

$$\Gamma, \psi \models^{l}_{\mathbf{V}\mathfrak{C}} \varphi \Leftrightarrow \Gamma \models^{l}_{\mathbf{V}\mathfrak{C}} \psi \supset \varphi$$

Proof. Straightforward by semantic conditions □

Global consequence has a peculiar deduction theorem:

Theorem 1.3 (Deduction Theorem)**.** *For all* $\Gamma \cup \{\varphi, \psi\} \subseteq For_{\mathcal{L}_{\Box\rightarrow}}$,

$$\Gamma, \psi \models^{g}_{\mathbf{V}\mathfrak{C}} \varphi \Leftrightarrow \text{there is a } n \in \mathbb{N} \text{ such that } \Gamma \models^{g}_{\mathbf{V}\mathfrak{C}} (\bigwedge_{m \leq n} \boxdot^{m}\psi) \supset \varphi$$

where $(\bigwedge_{m \leq n} \boxdot^{m}\psi)$ *is the formula* $\psi \wedge \boxdot\psi \wedge \boxdot^{2}\psi \wedge \cdots \wedge \boxdot^{n}\psi$

Proof.

(⇒) $\Gamma, \psi \models^{g}_{\mathbf{V}\mathfrak{C}} \varphi$. Then, by Theorem 1.1, we have that $\{\boxdot^{n}\gamma \mid \gamma \in \Gamma \text{ and } n \in \mathbb{N}\} \cup \{\boxdot^{n}\psi \mid n \in \mathbb{N}\} \models^{l}_{\mathbf{V}\mathfrak{C}} \varphi$. By definition of $\models^{l}_{\mathbf{V}\mathfrak{C}}$ and Remark 1.3, it is straightforward to see that there is a $m \in \mathbf{N}$ such that $\{\boxdot^{n}\gamma \mid \gamma \in \Gamma \text{ and } n \in \mathbb{N}\} \models^{l}_{\mathbf{V}\mathfrak{C}} (\bigwedge_{k \leq m} \boxdot^{m}\psi) \supset \varphi$. By Theorem 1.1, again, we then have that $\Gamma \models^{g}_{\mathbf{V}\mathfrak{C}} (\bigwedge_{k \leq m} \boxdot^{m}\psi) \supset \varphi$.

(⇐) Assume that there is a $n \in \mathbb{N}$ such that $\Gamma \models^{g}_{\mathbf{V}\mathfrak{C}} (\bigwedge_{m \leq n} \boxdot^{m}\psi) \supset \varphi$, namely for all spherical Lewisian models $\mathcal{M} = \langle W, \mathcal{S}, \models \rangle$ satisfying $\mathfrak{C}$, we have that if $\mathcal{M} \models \gamma$ for all $\gamma \in \Gamma$, then $\mathcal{M} \models (\bigwedge_{m \leq n} \boxdot^{m}\psi) \supset \varphi$. Now, consider any spherical Lewisian model $\mathcal{M}$ satisfying $\mathfrak{C}$ and assume $\mathcal{M} \models \gamma$ for all $\gamma \in \Gamma$ and $\mathcal{M} \models \psi$. By assumption, we have that for any word $w \in W$, $w \models (\bigwedge_{m \leq n} \boxdot^{m}\psi) \supset \varphi$. Moreover, by assumption we also have that for all $w \in W$, $w \models \psi$. Thus, a fortiori, by Lemma 1.3, we have that for all the worlds $w \in W$, $w \models \boxdot\gamma$ for all $\gamma \in \Gamma$. And for all $w \in W$, we have that $w \models \bigwedge_{m \leq n} \boxdot^{m}\psi$. Moreover, by assumption, we have that for all $w \in W$, $w \models (\bigwedge_{m \leq n} \boxdot^{m}\psi) \supset \varphi$. Hence, by semantic conditions of $\supset$ (i.e. semantic modus ponens), we have that for all $w \in W$, $w \models \varphi$ too. Hence $\Gamma, \psi \models^{g}_{\mathbf{V}\mathfrak{C}} \varphi$.

□

Now, we have all the ingredients to initiate an algebraic treatment of variably strict conditionals logics.

1.2 Algebraic Semantics

In this section, we will introduce an algebraic account of variably strict conditional logics. The results from the previous section will be essential to prove some key results concerning algebraizability of variably strict conditional logics. We assume the reader is familiar with basic notions of abstract algebraic logic and with the theory of Boolean algebras (see Burris and Sankappanavar 1981; Font 2016).

The fundamental structure of our approach is a kind of algebra that we refer to as *conditional algebra*.

Definition 1.8. *A conditional algebra is a tuple of the form* $\mathbf{V} = \langle V, \wedge, \vee, \neg, \Box\!\!\rightarrow, \top, \bot\rangle$ *where* $\langle V, \wedge, \vee, \neg, \bot, \top\rangle$ *is a Boolean algebra with* $\supset$ *and* $\leftrightarrow$ *defined as usual, and* $\Box\!\!\rightarrow$ *is a binary operation on* V *such that for all* $x, y, z \in V$*:*

1. $x \Box\!\!\rightarrow x = \top$
2. $((x \Box\!\!\rightarrow y) \wedge (y \Box\!\!\rightarrow x)) \wedge ((x \Box\!\!\rightarrow z) \leftrightarrow (y \Box\!\!\rightarrow z)) = (x \Box\!\!\rightarrow y) \wedge (y \Box\!\!\rightarrow x)$
3. $((x \vee y) \Box\!\!\rightarrow x) \vee ((x \vee y) \Box\!\!\rightarrow y) \vee (((x \vee y) \Box\!\!\rightarrow z) \leftrightarrow ((x \Box\!\!\rightarrow z) \wedge (y \Box\!\!\rightarrow z)) = \top$
4. $x \Box\!\!\rightarrow (y \wedge z) = (x \Box\!\!\rightarrow y) \wedge (x \Box\!\!\rightarrow z)$
5. $(x \Box\!\!\rightarrow (y \wedge z)) \supset (x \Box\!\!\rightarrow (y \vee z)) = \top$

Moreover, we define some non-primitive operations which intuitively correspond to the non-primitive connectives in the language $\mathcal{L}_{\Box\!\!\rightarrow}$*:*

- $(x \Diamond\!\!\rightarrow y) := \neg(x \Box\!\!\rightarrow \neg y)$*;*
- $\boxdot x := \neg x \Box\!\!\rightarrow x$*;*
- $\diamonddot x := \neg \boxdot \neg x$*;*
- $x \preccurlyeq y := ((x \vee y) \Diamond\!\!\rightarrow (x \vee y)) \supset ((x \vee y) \Diamond\!\!\rightarrow x)$

It is straightforward to observe that the family of conditional algebras is a variety since it can be characterized only by equations. Intuitively, the above equations correspond to the axioms and rules from the logic **V**. Specifically, equations 1, 2 and 3 corresponds respectively to axioms (ID), (CE), and (NE) from Definition 0.4; while 4 and 5 together correspond to the rule (DWC), as the next remark shows.

Remark 1.4. *Given a conditional algebra* $\mathbf{V} = \langle V, \wedge, \vee, \neg, \Box\!\!\rightarrow, \top, \bot\rangle$*, for* $x_1, \ldots, x_n, y, z \in V$*, we have that if* $(\bigwedge_{1 \leq i \leq n} x_i) \supset y = \top$*, i.e.* $(\bigwedge_{1 \leq i \leq n} x_i) \leq y$*, then* $(\bigwedge_{1 \leq i \leq n} z \Box\!\!\rightarrow x_i) \supset (z \Box\!\!\rightarrow y) = \top$

Proof. Assume $(x_1 \wedge x_2) \supset y = \top$. Then, by properties of Boolean algebras, we have that $x_1 \wedge x_2 \leq y$, thus $(x_1 \wedge x_2) \wedge y = (x_1 \wedge x_2)$ and $(x_1 \wedge x_2) \vee y = y$. By equation 5 in Definition 1.8, we have that $z \mathrel{\Box\!\!\rightarrow} (x_1 \wedge x_2) \supset z \mathrel{\Box\!\!\rightarrow} ((x_1 \wedge x_2) \vee y) = \top$. Therefore, $(z \mathrel{\Box\!\!\rightarrow} (x \wedge y)) \supset (z \mathrel{\Box\!\!\rightarrow} y) = \top$. By equation 4 in Definition 1.8, we have that $((z \mathrel{\Box\!\!\rightarrow} x_1) \wedge (z \mathrel{\Box\!\!\rightarrow} x_2)) \supset (z \mathrel{\Box\!\!\rightarrow} y) = \top$. We can reason analogously also with $x_1, \ldots x_n$.

□

Similarly, we can show that the operation $\mathrel{\Box\!\!\rightarrow}$ on a conditional algebra is monotone:

Remark 1.5. *Given a conditional algebra* $\mathbf{V} = \langle V, \wedge, \vee, \neg, \mathrel{\Box\!\!\rightarrow}, \top, \bot \rangle$, *for* $x, y \in V$, *if* $x \leq y$, *then* $z \mathrel{\Box\!\!\rightarrow} x \leq z \mathrel{\Box\!\!\rightarrow} y$. *Thus, if* $x = y$, *then* $z \mathrel{\Box\!\!\rightarrow} x = z \mathrel{\Box\!\!\rightarrow} y$

Analogously to the case of Lewisian models, we can define families of conditional algebras by imposing suitable equations:

	Equations on a conditional algebra
(N) Normality	$\boxdot x \supset \diamonddot x = \top$
(T) Total Reflexivity	$\boxdot x \supset x = \top$
(T) Weak Centering	$(x \mathrel{\Box\!\!\rightarrow} y) \supset (x \supset y) = \top$
(C) Centering	Weak Centering + $(x \wedge y) \supset (x \mathrel{\Box\!\!\rightarrow} y) = \top$
(S) Uniqueness	$(x \mathrel{\Box\!\!\rightarrow} y) \vee (x \mathrel{\Box\!\!\rightarrow} \neg y) = \top$
(U-) Local Uniformity	$\diamonddot x \supset \boxdot \diamonddot x = \top + \boxdot x \supset \boxdot \boxdot x = \top$
(U) Uniformity	$\diamonddot x \supset \boxdot \diamonddot x = \top + \boxdot x \supset \boxdot \boxdot x = \top$
(A-) Local Absoluteness	$(x \preccurlyeq y) \supset \boxdot (x \preccurlyeq y) = \top + (x \prec y) \supset \boxdot (x \preccurlyeq y) = \top$
(A-) Absoluteness	$(x \preccurlyeq y) \supset \boxdot (x \preccurlyeq y) = \top + (x \prec y) \supset \boxdot (x \preccurlyeq y) = \top$
(UT)Universality	**Uniformity + Total Reflexivity**
(WA) Weak Triviality	**Weak Centering + Absoluteness**
(CA) Triviality	**Centering + Absoluteness**

Table 1.1: A table summarizing the families of conditional algebras defined by the corresponding equation(s).

As we can observe from the table, we focus on the classes of conditional algebras that ideally correspond to variably strict conditional logics. Observe that the equations in Table 1.2 correspond to axioms in Table 3.1, with the only difference that they are written in the signature of conditional algebras.

Notation 1.3.

- *For an equation or a family of equations* $\mathfrak{C}$ *in Table 1.2,* $\mathbf{V}\mathfrak{C}$ *denotes the family of conditional algebras satisfying equation(s)* $\mathfrak{C}$.

The class $\mathbf{V\mathfrak{C}}$ of conditional algebras determines an algebraic semantic for the logic $\mathbf{V\mathfrak{C}}$. In particular, we are going to show that the global $\mathbf{V\mathfrak{C}}$-logics are algebraizable with respect to the corresponding $\mathbf{V\mathfrak{C}}$ class of conditional algebras, while the local $\mathbf{V\mathfrak{C}}$-logics are sound and complete with respect to a certain notion of logical consequence defined over conditional algebras. Firstly, we need some preliminaries.

Definition 1.9. *A valuation of our language on a conditional algebra* $\mathbf{V} = \langle V, \wedge, \vee, \neg, \Box\!\!\rightarrow, \top, \bot \rangle$ *is a homomorphism* $h : For_{\mathcal{L}_{\Box\!\rightarrow}} \to V$. *More explicitly:*

- *for all* $p \in Var$, $h(p) \in V$, *i.e.* h *maps* p *to an element of* V;
- *for all* $\varphi \in For_{\mathcal{L}_{\Box\!\rightarrow}}$, $h(\varphi)$ *is inductively defined as follows:*
  - $h(\neg\psi) = \neg h(\psi)$
  - $h(\psi \wedge \delta) = h(\psi) \wedge h(\delta)$;
  - $h(\psi \vee \delta) = h(\psi) \vee h(\delta)$;
  - $h(\psi \Box\!\!\rightarrow \delta) = h(\psi) \Box\!\!\rightarrow h(\delta)$;

Observe that the symbols $\neg, \wedge, \vee$, *and* $\Box\!\!\rightarrow$ *on the left side are the connectives from our language* $\mathcal{L}_{\Box\!\rightarrow}$, *while the symbols* $\neg, \wedge, \vee$, *and* $\Box\!\!\rightarrow$ *on th right side are the operations on the algebra* $\mathbf{V}$

The notion of equational (logical) consequence on a class of algebras K, with respect to our language $For_{\mathcal{L}_{\Box\!\rightarrow}}$ is defined as follows:

Definition 1.10. *For* $\{\varphi_i, \psi_i \mid i \in I\} \cup \{\delta, \epsilon\} \subseteq For_{\mathcal{L}_{\Box\!\rightarrow}}$,

$$\{\varphi_i = \psi_i \mid i \in I\} \models_K \delta = \epsilon \quad \Leftrightarrow \quad \begin{array}{l} \textit{for all algebras } \mathbf{A} \textit{ in the class } K, \\ \textit{for all valuations } h : For_{\mathcal{L}_{\Box\!\rightarrow}} \to \mathbf{A}, \\ \textit{if } h(\varphi_i) = h(\psi_i) \textit{ for all } i \in I, \textit{ then } h(\delta) = h(\epsilon) \end{array}$$

We define a notion of equational logical consequence over a class of conditional algebras as follows:

Definition 1.11. *For* $\{\varphi_i, \psi_i \mid i \in I\} \cup \{\delta, \epsilon\} \subseteq For_{\mathcal{L}_{\Box\!\rightarrow}}$,

$$\{\varphi_i = \psi_i \mid i \in I\} \models_{\mathbf{V\mathfrak{C}}} \delta = \epsilon \quad \Leftrightarrow \quad \begin{array}{l} \textit{for all algebras } \mathbf{V} \textit{ in the class } \mathbf{V\mathfrak{C}}, \\ \textit{for all valuations } h : For_{\mathcal{L}_{\Box\!\rightarrow}} \to \mathbf{V}, \\ \textit{if } h(\varphi_i) = h(\psi_i) \textit{ for all } i \in I, \textit{ then } h(\delta) = h(\epsilon) \end{array}$$

We focus on a special relation of equational consequence that is defined as follows: for all $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}_{\Box\!\rightarrow}}$,

$$\{\gamma = \top \mid \gamma \in \Gamma\} \models_{\mathbf{V}\mathfrak{C}} \varphi = \top \quad \Leftrightarrow \quad \begin{array}{l} \textit{for all algebras } \mathbf{V} \textit{ in the family } \mathbf{V}\mathfrak{C}, \\ \textit{for al valuations } h : For_{\mathcal{L}_{\Box\to}} \to \mathbf{V}, \\ \textit{if } h(\gamma) = \top \textit{ for all } \gamma \in \Gamma, \textit{ then } h(\varphi) = 1 \end{array}$$

The notion of equational logical consequence is sound and complete with respect to the corresponding global $\mathbf{V}\mathfrak{C}$-logic.

Theorem 1.4. *For* $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}_{\Box\to}}$,

$$\Gamma \vdash^{g}_{\mathbf{V}\mathfrak{C}} \varphi \Leftrightarrow \{\gamma = \top \mid \gamma \in \Gamma\} \models_{\mathbf{V}\mathfrak{C}} \varphi = \top$$

Proof.

($\Rightarrow$) First, we show that for all substitutional instances of the axioms and rules for global $\mathbf{V}\mathfrak{C}$ of the form $\Gamma \vdash^{g}_{\mathbf{V}\mathfrak{C}} \varphi$, we have that $\{\gamma = \top \mid \gamma \in \Gamma\} \models_{\mathbf{V}\mathfrak{C}} \varphi = \top$. Specifically, for all conditional algebras $\mathbf{V} = \langle V, \wedge, \vee, \neg, \Box\!\!\to, \top, \bot \rangle$, for all valuations $h : For_{\mathcal{L}_{\Box\to}} \to V$, the following hold:

- ($TAUT$) if φ is a classical tautology then $h(\varphi) = \top$, by properties of Boolean algebras;
- (ID) by equation 1 in Definition 1.8;
- (CE) by equation 2 in Definition 1.8;
- (NE) by equation 3 in Definition 1.8;
- * ($MIGHT$), (BOX), ($DIAM$), ($PLUS$) follow from definitions of non primitive operations in Definition 1.8;
- (MP) follows from properties of Boolean algebras, in particular the following consequence holds: $\varphi = \top, \varphi \supset \psi = 1 \models_{\mathbf{V}\mathfrak{C}} \psi = 1$;
- ($RTAUT1$) follows from properties of Boolean algebras;
- ($RTAUT2$) follows from properties of Boolean algebras;
- ($DWCG$) by Remark 1.4;
- (SUB) is a consequence of the following property; in any conditional algebras, if $x \leftrightarrow y = 1$, then $x = y$, so x and y can be substituted with each other;
- ($\mathfrak{C}$) follows from equation $\mathfrak{C}$ (see Table 1.2).

Now, assume that $\Gamma \vdash^{g}_{\mathbf{V}\mathfrak{C}} \varphi$, then there is a proof of φ from Γ, i.e. there is a finite sequence $\langle \varphi_1, \ldots, \varphi_n, \rangle$ of formulas with $\varphi_n = \varphi$ such that for all φ_i (for $1 \leq i \leq n$), either $\varphi_i \in \Gamma$, or φ_i and is an instantiation of an axiom schema, or it results from an application of a rule. We reason by strong induction on n. Now we have three cases to consider:

(i) φ_n is an instantiation of an axiom schema. Then, by what we proved above we have that $\models_{\mathbf{V}\mathfrak{C}} \varphi_n = 1$ since φ_n is an axiom;

(ii) $\varphi_n \in \Gamma$. Then, obviously $\{\gamma = 1 : \gamma \in \Gamma\} \models_{\mathbf{V}\mathfrak{C}} \varphi_n = 1$ since φ_n is already in Γ;

(iii) φ_n result from an application of a rule $\Delta \vdash^g_{\mathbf{V}\mathfrak{C}} \varphi_n$ such that all the formulas in Δ occur in the sequence $\langle \varphi_1, \ldots, \varphi_{n-1} \rangle$. Then, by induction hypothesis, it is the case that $\{\gamma = 1 : \gamma \in \Gamma\} \models_{\mathbf{V}\mathfrak{C}} \delta = 1$ for all $\delta \in \Delta$, and moreover $\{\delta = 1 : \delta \in \Delta\} \models_{\mathbf{V}\mathfrak{C}} \varphi_n = 1$. Then, by transitivity of $\models_{\mathbf{V}\mathfrak{C}}$, we have $\{\gamma = 1 : \gamma \in \Gamma\} \models_{\mathbf{V}\mathfrak{C}} \varphi_n = 1$.

($\Leftarrow$) For the other direction, we reason by contrapositio. Suppose $\Gamma \nvdash^g_{\mathbf{V}\mathfrak{C}} \varphi$. First, we prove that the relation $\theta \subseteq For_{\mathcal{L}_{\Box\!\rightarrow}} \times For_{\mathcal{L}_{\Box\!\rightarrow}}$ defined below is a congruence relation:

$$\theta := \{(\varphi, \psi) \in For_{\mathcal{L}_{\Box\!\rightarrow}} \times For_{\mathcal{L}_{\Box\!\rightarrow}} : \Gamma \vdash^g_{\mathbf{V}\mathfrak{C}} (\varphi \supset \psi) \wedge (\psi \supset \varphi)\}$$

Namely, for all the connectives, we have to show that:

1. if $\varphi\theta\psi$ and, then $\neg\varphi\theta\neg\epsilon$
2. if $\varphi\theta\psi$ and $\gamma\theta\epsilon$, then $\varphi \wedge \gamma\theta\psi \wedge \epsilon$
3. if $\varphi\theta\psi$ and $\gamma\theta\epsilon$, then $\varphi \vee \gamma\theta\psi \vee \epsilon$
4. if $\varphi\theta\psi$ and $\gamma\theta\epsilon$, then $\varphi \Box\!\!\rightarrow \gamma\theta\psi \Box\!\!\rightarrow \epsilon$

1, 2 and 3 are straightforward and follows from the fact that every conditional algebra satisfies the equations of Boolean algebras. The only interesting case is te counterfactual connective: assume $\varphi\theta\psi$ and $\delta\theta\epsilon$, namely $\Gamma \vdash^g_{\mathbf{V}\mathfrak{C}} \varphi \leftrightarrow \psi$. Then, by $(DWCG)$ and $(SUBG)$, we can easily derive that $\Gamma \vdash^g_{\mathbf{V}\mathfrak{C}} (\varphi \Box\!\!\rightarrow \delta) \leftrightarrow (\psi \Box\!\!\rightarrow \epsilon)$. Hence, $(\varphi \Box\!\!\rightarrow \delta)\theta(\psi \Box\!\!\rightarrow \epsilon)$

Now, consider the Lindenbaum algebra of the logic $\vdash^g_{\mathbf{V}\mathfrak{C}}$ with respect to the congruence θ, i.e. $For^{/\theta}_{\mathcal{L}_{\Box\!\rightarrow}}$. It is the algebras $\langle For^{/\theta}_{\mathcal{L}_{\Box\!\rightarrow}}, \wedge, \vee, \neg, \Box\!\!\rightarrow [\top]_\theta, [\bot]_\theta \rangle$ where:

- $For^{/\theta}_{\mathcal{L}_{\Box\!\rightarrow}} = \{[\varphi]_\theta \mid \varphi \in For_{\mathcal{L}_{\Box\!\rightarrow}}\}$ is the set of equivalence classes of all the formulas in $For_{\mathcal{L}_{\Box\!\rightarrow}}$ modulo θ.
- $\wedge$ is defined as: $[\varphi]_\theta \wedge [\psi]_\theta = [\varphi \wedge \psi]_\theta$
- $\vee$ is defined as: $[\varphi]_\theta \vee [\psi]_\theta = [\varphi \vee \psi]_\theta$
- $\neg$ is defined as: $\neg[\varphi]_\theta = [\neg\varphi]_\theta$
- $\Box\!\!\rightarrow$ is defined as: $[\varphi]_\theta \Box\!\!\rightarrow [\psi]_\theta = [\varphi \Box\!\!\rightarrow \psi]_\theta$

By the valid inferences of of $\vdash^{g}_{\mathbf{V}\mathfrak{C}}$, it is easy to show that $\langle For^{/\theta}_{\mathcal{L}_{\Box\to}}, \wedge, \vee, \neg, \Box\!\!\to [\top]_\theta, [\bot]_\theta \rangle$ is indeed a conditional algebra satisfying axiom $\mathfrak{C}$ where $[\top]_\theta$ is its top element.

Now, consider the valuation $h : For_{\mathcal{L}_{\Box\to}} \to For^{/\theta}_{\mathcal{L}_{\Box\to}}$ such that:

- for all $\varphi \in For_{\mathcal{L}_{\Box\to}}$, $h(\varphi) = [\varphi]_\theta$

It is easy to show that h is indeed a homomorphism. Now, observe that for all $\gamma \in \Gamma$, $h(\gamma) = \top$. Indeed, by definition of θ, we have that for all $\gamma \in \Gamma$, $\Gamma \vdash^{g}_{\mathbf{V}\mathfrak{C}} (\gamma \supset \top) \wedge (\top \supset \gamma)$. Hence, it is the case that $h(\gamma) = [\top]_\theta$. Recall that our assumption was $\Gamma \nvdash^{g}_{\mathbf{V}\mathfrak{C}} \varphi$. Now, suppose $h(\varphi) = [\top]_\theta$. By definition of θ, this implies that $\Gamma \vdash^{g}_{\mathbf{V}\mathfrak{C}} (\varphi \supset \top) \wedge (\top \supset \varphi)$, and so that $\Gamma \vdash^{g}_{\mathbf{V}\mathfrak{C}} \varphi$, which is in contradiction with our assumption. Hence, it must be the case that $h(\varphi) \neq [\top]_\theta$. And this implies $\{\gamma = \top \mid \gamma \in \Gamma\} \not\models_{\mathbf{V}\mathfrak{C}} \varphi = 1$

□

We have proven that the global $\mathbf{V}\mathfrak{C}$-logic is sound and complete with respect to the equational logical consequence over the corresponding conditional algebras in the family $\mathbf{V}\mathfrak{C}$. However, the completeness result is just an initial step in the direction of proving algebraizability for variably strict conditionals logics. And this will be the topic of the next subsection.

However, before moving on to the exploration of algebraizability, we will show that the local $\mathbf{V}\mathfrak{C}$-logic does not behave analogously to its global counterpart. In particular, we need a different notion of logical consequence to characterize the local $\mathbf{V}\mathfrak{C}$ logics over conditional algebras:

Definition 1.12. *The relation of preserving degrees of truth on conditional algebras is defined as follows: for all $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}_{\Box\to}}$,*

$$\Gamma \models^{\leq}_{\mathbf{V}\mathfrak{C}} \varphi \quad \Leftrightarrow \quad \begin{array}{l}\textit{for all conditional algebras } \mathbf{V} \textit{ in the family } \mathbf{V}\mathfrak{C},\\ \textit{for all valuations } h : For_{\mathcal{L}_{\Box\to}} \to V, \textit{ and all } a \in V,\\ \textit{if } a \leq h(\gamma) \textit{ for all } \gamma \in \Gamma, \textit{ then } a \leq h(\varphi)\end{array}$$

The relation of preserving degrees of truth over the family $\mathbf{V}\mathfrak{C}$ of conditional algebras is sound and complete with respect to the corresponding local $\mathbf{V}\mathfrak{C}$-logic:

Theorem 1.5. *For all $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}_{\Box\to}}$,*

$$\Gamma \vdash^{l}_{\mathbf{V}\mathfrak{C}} \varphi \Leftrightarrow \Gamma \models^{\leq}_{\mathbf{V}\mathfrak{C}} \varphi$$

Proof.

(⇒) This direction proceeds similarly to (⇒) in the proof of Theorem 1.4

(⇐) We reason by contraposition. Assume $\Gamma \nvdash^{\leq}_{\mathbf{V}\mathfrak{C}} \varphi$. Then consider the relation $\theta \subseteq For_{\mathcal{L}_{\Box\!\!\rightarrow}} \times For_{\mathcal{L}_{\Box\!\!\rightarrow}}$ defined as follows:

$$\theta := \{(\varphi, \psi) \in For_{\mathcal{L}_{\Box\!\!\rightarrow}} \times For_{\mathcal{L}_{\Box\!\!\rightarrow}}\} : \Gamma \vdash^{l}_{\mathbf{V}\mathfrak{C}} \boxdot^{n}(\varphi \supset \psi) \text{ and } \Gamma \vdash^{l}_{\mathbf{V}\mathfrak{C}} \boxdot^{n}(\psi \supset \varphi) \text{ for all } n \in \mathbb{N}\,\}$$

It is not difficult to show that θ is a congruence relation. For exemplification, we will show that θ preserve the binary operation $\Box\!\!\rightarrow$. Assume $\varphi\theta\psi$ and $\delta\theta\epsilon$. Hence we have that:

(1) $\Gamma \vdash^{l}_{\mathbf{V}\mathfrak{C}} \boxdot^{n}(\varphi \supset \psi) \wedge \boxdot^{n}(\psi \supset \varphi)$ for all $n \in \mathbb{N}$

notice that, from (1), it follows that for any $n \in \mathbb{N}$,

(2) $\Gamma \vdash^{l}_{\mathbf{V}\mathfrak{C}} \boxdot^{n+1}(\varphi \supset \psi) \wedge \boxdot^{n+1}(\psi \supset \varphi)$

from (1) and (2), we can infer that

(3) $\Gamma \vdash^{l}_{\mathbf{V}\mathfrak{C}} \boxdot^{n}((\epsilon \Box\!\!\rightarrow \varphi)) \supset (\epsilon \Box\!\!\rightarrow \psi))$ for all $n \in \mathbb{N}$

(4) $\Gamma \vdash^{l}_{\mathbf{V}\mathfrak{C}} \boxdot^{n}((\epsilon \Box\!\!\rightarrow \psi) \supset (\epsilon \Box\!\!\rightarrow \varphi))$ for all $n \in \mathbb{N}$

Hence, we conclude that $(\epsilon \Box\!\!\rightarrow \psi, \epsilon \Box\!\!\rightarrow \varphi) \in \theta$. Since we assumed that $\epsilon\theta\delta$, we can reason similarly and show that $(\epsilon \Box\!\!\rightarrow \psi, \delta \Box\!\!\rightarrow \varphi) \in \theta$. And analogously, $(\varphi \Box\!\!\rightarrow \epsilon, \psi \Box\!\!\rightarrow \delta) \in \theta$.

The above reasoning deserves some additional comments. Indeed, consider $\boxdot(\varphi \supset \psi) \wedge \boxdot(\psi \supset \varphi)$. By definition of $\boxdot$, this is equivalent to

$$(\neg(\varphi \supset \psi) \Box\!\!\rightarrow (\varphi \supset \psi)) \wedge (\neg(\psi \supset \varphi) \Box\!\!\rightarrow (\psi \supset \varphi))$$

Let us examine this formula from a semantic perspective. This formula is true at a world w in spherical Lewisian model if and only if all the worlds w' appearing in $\mathcal{S}(w)$ makes either φ and ψ both true, or both false. Hence, it is straightforward to see that also $(\epsilon \Box\!\!\rightarrow \varphi) \supset (\epsilon \Box\!\!\rightarrow \psi)$ and $(\epsilon \Box\!\!\rightarrow \psi) \supset (\epsilon \Box\!\!\rightarrow \varphi)$ must be true at w.

Another comment would be illuminating. Consider an alternative relation

$$\theta' := \{(\varphi, \psi) \in For_{\mathcal{L}_{\Box\!\!\rightarrow}} \times For_{\mathcal{L}_{\Box\!\!\rightarrow}} : \Gamma \vdash^{l}_{\mathbf{V}\mathfrak{C}} (\varphi \supset \psi) \wedge (\psi \supset \varphi)\}$$

which is analogous to the one in the proof of Theorem 1.4. This relation is not a congruence since it does not preserve the counterfactual connective. Indeed, from a semantic perspective, it might be the case that $x \leftrightarrow y$ is true at a world w in a spherical Lewisian model, but $(z \Box\!\!\rightarrow x) \supset (z \Box\!\!\rightarrow y)$ is not true at w.

Now, consider the Lindenbaum algebra of the logic $\vdash^{l}_{\mathbf{V}\mathfrak{C}}$ with respect to the congruence θ, i.e. $For^{/\theta}_{\mathcal{L}_{\Box\!\rightarrow}}$. It is the algebras $\langle For^{/\theta}_{\mathcal{L}_{\Box\!\rightarrow}}, \wedge, \vee, \neg, \Box\!\!\rightarrow [\top]_{\theta}, [\bot]_{\theta}\rangle$ defined as in the proof of Theorem 1.4.

Now, to prove our main claim, we observe that

$$\Gamma \vdash^{l}_{\mathbf{V}\mathfrak{C}} \varphi \Leftrightarrow \text{ there is some finite } \Delta \subseteq \Gamma \text{ such that } \Delta \vdash^{l}_{\mathbf{V}\mathfrak{C}} \varphi$$

Consider the canonical valuation $h : For_{\mathcal{L}_{\Box\!\rightarrow}} \rightarrow For^{/\theta}_{\mathcal{L}_{\Box\!\rightarrow}}$ and suppose that there is some finite $\Delta = \{\delta_1, \dots, \delta_n\} \subseteq \Gamma$ such that $h(\delta_1) \wedge \dots \wedge h(\delta_n) \leq h(\varphi)$. Since h is an homomorphism, we obtain that $h((\delta_1 \wedge \dots \delta_n) \vee \varphi) = h(\varphi)$; hence, we get that $\Gamma \vdash^{l}_{\mathbf{V}\mathfrak{C}} ((\delta_1 \wedge \dots \wedge \delta_n) \vee \varphi) \supset \varphi$. Clearly we have that $\Gamma \vdash^{l}_{\mathbf{V}\mathfrak{C}} \delta_1 \wedge \dots \wedge \delta_n$ and so $\Gamma \vdash^{l}_{\mathbf{V}\mathfrak{C}} (\delta_1 \wedge \dots \wedge \delta_n) \supset \varphi$. Then, by modus ponens, $\Gamma \vdash^{l}_{\mathbf{V}\mathfrak{C}} \varphi$, which is in contradiction with our assumption. Hence, it must the case that $h(\delta_1) \wedge \dots \wedge h(\delta_n) \nleq h(\varphi)$. Therefore, $\Gamma \not\models^{\leq}_{\mathbf{V}\mathfrak{C}} \varphi$.

□

As a corollary of the above theorem, we get that:

Corollary 1.1. *For all $\varphi, \psi \in For_{\mathcal{L}_{\Box\!\rightarrow}}$,*

$$\models_{\mathbf{V}\mathfrak{C}} \varphi = \psi \Leftrightarrow \varphi \dashv\vdash^{l}_{\mathbf{V}\mathfrak{C}} \psi$$

In the following, we initiate a deeper algebraic investigation of local and global $\mathbf{V}\mathfrak{C}$-logics.

1.2.1 Algebraizability

The notion of algebraizability of a logic is due to Blok and Pigozzi (2014). The idea is that algebraic structures and a logic must be connected in a very profound way that goes beyond the simple soundness and completeness results. Namely, there must exist a tight connection between deductions within a logic and equational consequences within a family of algebras. The following definition will clarify the concept:

Notation 1.4.

- *$E(x)$ indicates a set of equations where just one variable $x \in Var$ appears. Possible examples might be:*
  - $E(x) = \{x = \top\}$
  - $E(x) = \{x = \neg x\}$

- $E(x) = \{x = x \wedge x\}$

- *$F(x, y)$ indicates a set of formulas where only two variables $x, y \in Var$ appear. Possible examples might be:*
  - $F(x, y) = \{x \supset y, y \supset x\}$
  - $F(x, y) = \{x \supset \neg y, y \supset x\}$
  - $F(x, y) = \{x \,\square\!\!\rightarrow y, y \,\square\!\!\rightarrow x\}$

- *Given a set of formulas and a set of equations $E(x)$, $E(\{\varphi_1, \ldots, \varphi_n\})$ is the set of all and only the equations eq obtained as follows: for each equation eq′ in $E(x)$, eq is obtained by replacing x in eq with some φ_i. Some examples:*
  - *For $E(x) = \{x = 1\}$, and $F(x, y) = \{x \supset y, y \supset x\}$, $E(F(x, y)) = \{x \supset y = 1, y \supset x = 1\}$*
  - *For $E(x) = \{x = \neg x\}$, and $F(x, y) = \{x \wedge y\}$, $E(F(x, y)) = \{x \wedge y = \neg(x \wedge y)\}$*

Definition 1.13 (Blok and Pigozzi 2014; Font 2016). *A logic $\vdash$ is algebraizable if there is a class of algebras K, a set of equations $E(x)$ and a set of formulas $F(x, y)$ such that:*

- $\Gamma \vdash \varphi \Leftrightarrow \{E(\gamma) \mid \gamma \in \Gamma\} \models_K E(\varphi)$;
- *for all equations eq $\in E(F(x, y))$, $x = y \models_K eq$;*
- $E(F(x, y)) \models_K x = y$

Now, we are going to show that the global $\mathbf{V}\mathfrak{C}$-logics are algebraizable with respect to the corresponding family $\mathbf{V}\mathfrak{C}$ of conditional algebras:

Theorem 1.6. *There is a set of equations in one variable $E(x)$ and a set of formulas $F(x, y)$ such that the following hold:*

- *for all equations eq $\in E(F(x, y))$, $x = y \models_{\mathbf{V}\mathfrak{C}} eq$*
- $E(F(x, y)) \models_{\mathbf{V}\mathfrak{C}} x = y$

Proof. Consider the sets:

$$E(x) = \{x = \top\} \text{ and } F(x, y) = \{x \supset y, y \supset x\}$$

Notice that

$$x = y \models_{\mathbf{V}\mathfrak{C}} (x \supset y) \wedge (y \supset x) = \top \text{ and } (x \supset y) \wedge (y \supset x) = \top \models_{\mathbf{V}\mathfrak{C}} x = y$$

since for any element a, b in a conditional algebra the following holds:

$$\begin{aligned} a = b \quad &\Leftrightarrow \quad a \leq b \text{ and } b \leq a \\ &\Leftrightarrow \quad a \supset b = \top \text{ and } b \supset a = \top \end{aligned}$$

The last step of the above argument is given by the fact that every conditional algebra is a Boolean algebra. □

The above observation establishes the desired result of having a tight connection between deductions within a logic and equational consequences over an algebra. Indeed, by combining Theorem 1.4 and Theorem 1.6, we obtain the following corollary:

Corollary 1.2. *Any global* $\mathbf{V}\mathfrak{C}$*-logic is algebraizable and its algebraizability is witnessed by the set of equations* $E(x) = \{x = \top\}$*, the set of formulas* $F(x, y) = \{x \supset y, y \supset x\}$*, and the variety of conditional algebras* $\mathbf{V}\mathfrak{C}$*.*

In other words, the conditions established by Theorem 1.4 and Theorem 1.6 are necessary and sufficient conditions to make the global variably strict conditional logics algebraizable, according to the the general theory of abstract algebraic logic (see for instance Font 2016; Moraschini 2023).

At this point, a natural question arises regarding the algebraic properties of local $\mathbf{V}\mathfrak{C}$-logics. However, in order to investigate the algebraizability of the local $\mathbf{V}\mathfrak{C}$-logics, it is essential to introduce some new definitions and recap some general results.

Bits of Structure Theory

In what follows, we provide a characterization of the deductive filters of local $\mathbf{V}\mathfrak{C}$ logics. Firstly, we introduce the definition of a deductive filter.

Definition 1.14. *Given an conditional algebra* $\mathbf{V} = \langle V, \wedge, \vee, \neg, \Box\!\!\rightarrow, \top, \bot \rangle$*, a deductive filter of the logic* $\vdash^{l}_{\mathbf{V}\mathfrak{C}}$ *on* $\mathbf{V}$ *is a subset* $F \subseteq V$ *of* V *such that: for all* $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}_{\Box\rightarrow}}$*,*

$$\begin{aligned} \Gamma \vdash^{l}_{\mathbf{V}\mathfrak{C}} \varphi \quad \Leftrightarrow \quad & \textit{for all valuations } h : For_{\mathcal{L}_{\Box\rightarrow}} \to V, \\ & \textit{if } h(\gamma) \in F \textit{ for all } \gamma \in \Gamma, \textit{ then } h(\varphi) \in F \end{aligned}$$

An additional definition would be useful:

Definition 1.15. *Given a conditional algebra* $\mathbf{V} = \langle V, \wedge, \vee, \neg, \Box\!\!\rightarrow, \top, \bot \rangle$*, a lattice filter on* $\mathbf{V}$ *is a subset* $F \subseteq V$ *of* V *such that: for all* $a, b \in V$*, the following hold:*

1. $\top \in F$
2. *if* $a, b \in F$*, then* $a \wedge b \in F$*;*

3. *if $a \in F$ and $a \leq b$, then $b \in F$*

Now, we have all the ingredients to provide an informative characterization of the deductive filters of local $\mathbf{V}\mathfrak{C}$-logics:

Lemma 1.5. *Given a conditional algebra $\mathbf{V} = \langle V, \wedge, \vee, \neg, \Box\!\!\rightarrow, \top, \bot \rangle$, the following condition holds:*

$$\text{a subset } F \subseteq V \text{ is a deductive filter of } \vdash^{l}_{\mathbf{V}\mathfrak{C}} \text{ on } \mathbf{V} \Leftrightarrow F \text{ is a lattice filter on } \mathbf{V}$$

Proof. Mutatis mutandis, the proof is similar to the direction ($\Rightarrow$) of the proof of Theorem 1.4. We will show one case for exemplification, in particular that lattice filters preserve modus ponens. Consider $\varphi, \varphi \supset \psi \vdash^{l}_{\mathbf{V}\mathfrak{C}} \psi$, any conditional algebra $\mathbf{V}$, and any valuation $h : For_{\mathcal{L}_{\Box\!\rightarrow}} \to V$. Consider a lattice filter $F \subseteq V$ and suppose $h(\varphi) \in F$ and $h(\varphi \supset \psi) \in F$. Then, by definition of lattice filter, we have that $h(\varphi) \wedge h(\varphi \supset \psi) \in F$. Since h is an homomorphism, we have that $h(\varphi) \wedge (h(\varphi) \supset h(\psi))$. Since $h(\varphi) \wedge (h(\varphi) \supset h(\psi)) = h(\psi)$. Therefore $h(\psi) \in F$. □

Bits of Duality

In this subsection, we will show how a special finite spherical Lewisian model satisfying condition(s) $\mathfrak{C}$ induces a conditional algebra satisfying the corresponding equation $\mathfrak{C}$, and vice versa.

Lemma 1.6. *Consider a finite spherical Lewisian model $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$ where:*

- $W = \{w_1, w_2\}$
- *$\mathcal{S}$ is such that:*
  - $\mathcal{S}(w_1) = \{\{w_1\}, \{w_1, w_2\}\}$
  - $\mathcal{S}(w_2) = \{\{w_2\}, \{w_1, w_2\}\}$
- *$\vDash$ is such that:*
  - *$w_1 \vDash p$ and $w_2 \nvDash p$*
  - *$w_1 \nvDash q$ and $w_2 \nvDash q$*

$\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$ satisfies conditions **(N)**, **(T)**, **(W)**, **(C)**, **(S)**, **(U)**, **(U-)**, **(UT)**.

Then, the structure $\langle \wp(W), \cap, \cup, ^{-}, \Box\!\!\rightarrow, W, \emptyset \rangle$ is conditional algebra satisfying equations **(N)**, **(T)**, **(W)**, **(C)**, **(S)**, **(U)**, **(U-)**, **(UT)**, *in which*

- *$\cap$ is set-theoretic intersection*

- $\cup$ *is set-theoretic union*
- $^-$ *is set-theoretic complement*
- $\Box\!\!\rightarrow : \wp(W) \times \wp(W) \to \wp(X)$ *is defined as follows: for all* $X, Y \in \wp(W)$,

$$X \Box\!\!\rightarrow Y = \{w \in W \mid min^X_{\subseteq}(\mathcal{S}(w)) \subseteq Y\}$$

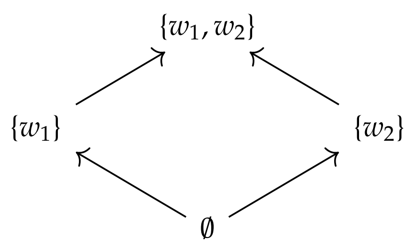
where

$$min^X_{\subseteq}(\mathcal{S}(w)) = \begin{cases} \textit{the minimal } S \in \mathcal{S}(w) \textit{ such that } S \cap X \neq \emptyset & \textit{if } \bigcup \mathcal{S}(w) \cap X \neq \emptyset \\ \emptyset & \textit{otherwise} \end{cases}$$

Moreover, it holds that

- $\boxdot\{w_1\} = \boxdot\{w_2\} = \boxdot\emptyset = \emptyset$,
- $\boxdot W = W$

Proof. The algebra can be depicted as:



$$\boxdot\{w_1, w_2\} = \{w_1, w_2\}$$

$$\boxdot\{w_1\} = \boxdot\{w_2\} = \boxdot\emptyset = \emptyset$$

It is easy to see that $\langle \wp(W), \cap, \cup, ^-, W, \emptyset \rangle$ is a Boolean algebra where $\cap$ is the meet operation, $\cup$ the join, W and $\emptyset$ are the top and the bottom element respectively. Moreover, observe that $[p] = \{w_1\}$, $[\neg p] = \{w_2\}$, $[p \vee \neg p] = \{w_1, w_2\}$ and $[q] = \emptyset$. For, for the sake of simplicity, we can identify the elements of $\wp(W)$ with $[q], [p], [\neg p], [p \vee \neg p]$.

Additionally, observe that $\mathcal{S}$ satisfies the constraints **(N)**, **(T)**, **(W)**, **(C)**, **(S)**, **(U)**, **(U-)**, **(UT)**. Hence, it must be the case that the corresponding axiom for each of those constraints are true at all the worlds in W. This implies that the operation of $\Box\!\!\rightarrow$ in $\langle \wp(W), \cap, \cup, ^-, \Box\!\!\rightarrow, W, \emptyset \rangle$ satisfies the corresponding equations in Table 1.2 and the equations characterizing conditional algebras in Definition 1.8. We will show one case for exemplification, namely that the equation **(W)** holds in $\langle \wp(W), \cap, \cup, ^-, W, \emptyset \rangle$, that is, for all $X, Y \in \wp(W)$, $X \Box\!\!\rightarrow Y \subseteq X^- \cup Y$. We have different cases to consider:

- if $X = \emptyset$, then clearly $X \boxright Y \subseteq X^- \cup Y$ since $X^- = W$.
- if $Y = W$, the we reason similarly to the case above.
- if $Y = \emptyset$, then, unless $X = Y$, we have that $X \boxright Y = \emptyset$. So **(W)** clearly holds.
- if $X = W$, then we have that $min^X_\subseteq(\mathcal{S}(w_1)) = \{w_1\}$ and $min^X_\subseteq(\mathcal{S}(w_2)) = \{w_2\}$. Hence:
  - if $Y = \{w_1\}$, this means that $X \boxright Y = \{w_1\}$. Clearly $X^- \cup Y = \{w_1\}$, and so **(W)** holds
  - if $Y = \{w_2\}$, this means that $X \boxright Y = \{w_2\}$. Clearly $X^- \cup Y = \{w_2\}$, and so **(W)** holds
  - if $Y = \emptyset$, then $X \boxright Y = \emptyset$ and so clearly **(W)** holds.
- if $X = \{w_1\}$ and $Y = \{w_2\}$, this means that $X \boxright Y = [p \boxright \neg p] = \emptyset$. And so **(W)** holds. Analogously if $X = \{w_2\}$ and $Y = \{w_1\}$

So, the main idea is that we can rely on the semantic conditions of $\boxright$ in $\mathcal{M}$ to show that the operation $\boxright$ satisfies the relevant equations. The key observation is that for all $X \boxright Y$, we have that either $X \boxright Y = [p \boxright \neg p]$, or $X \boxright Y = [(p \wedge \neg p) \boxright p]$, or $X \boxright Y = [(p \vee \neg p) \boxright p]$ or $X \boxright Y = [(p \vee \neg p) \boxright \neg p]$.

Moreover, observe that $\boxdot\{w_1\} = [\neg p \boxright p] = \emptyset$; $\boxdot\{w_2\} = [p \boxright \neg p] = \emptyset$; $\boxdot\emptyset = [\neg q \boxright q] = \emptyset$; $\boxdot W = [(p \wedge \neg p) \boxright (p \vee \neg p)] = W$

□

All the observations concerning deductive filters and the special examples of conditional algebras we have provided are essential to prove that most of local $\mathbf{V}\mathfrak{C}$-logics are not algebrizable, specifically those local $\mathbf{V}\mathfrak{C}$-logics such that $\mathfrak{C}$ is an axiom (or a family of axioms) among those in $\{$**(N)**, **(T)**, **(W)**, **(C)**, **(S)**, **(U)**, **(U-)**, **(UT)**$\}$. For this purpose, we begin by showing that if any of such logics were algebraizable, then they would be algebraizable with respect to a class K of conditional algebras.

Lemma 1.7. *Let $\mathfrak{C}$ be a condition or a family of conditions among those in* $\{$**(N)**, **(T)**, **(W)**, **(C)**, **(S)**, **(U)**, **(U-)**, **(UT)**$\}$. *If the local $\mathbf{V}\mathfrak{C}$-logic were algebraizable, then the class of algebras witnessing the algebraizability of $\mathbf{V}\mathfrak{C}$, according to Definition 1.13, must be a class of conditional algebras satisfying $\mathfrak{C}$.*

Proof. Suppose local $\mathbf{V}\mathfrak{C}$ is algebraizable. Then, according to Definition 1.13, there must be a set of equations $E(x)$, a set of formulas $F(x, y)$, and a class of algebras K that witness the algebraizbaility of $\mathbf{V}\mathfrak{C}$. We are going to show that K must be a class of conditional algebras satisfying $\mathfrak{C}$. For this purpose, it suffices to show that every equation that is valid in the conditional algebras in $\mathbf{V}\mathfrak{C}$ is also valid in all the algebras in K. Assume $\varphi = \psi$ is an equation that holds in the class of conditional algebras $\mathbf{V}\mathfrak{C}$, namely all the algebras in the class $\mathbf{V}\mathfrak{C}$ satisfy $\varphi = \psi$. By assumption, the local $\mathbf{V}\mathfrak{C}$-logic is algebraizable, hence it should be the case that

$$\vdash^{l}_{\mathbf{V}\mathfrak{C}} F(x, x) \Leftrightarrow \models_K x = x$$

So, by uniform substitution,

$$\vdash^{l}_{\mathbf{V}\mathfrak{C}} F(\varphi, \varphi) \Leftrightarrow \models_K \varphi = \varphi$$

Moreover, observe that the that the following holds:

$$\models_K \varphi = \varphi$$

Thus, by algebraizability of local $\mathbf{V}\mathfrak{C}$, it must hold that $\vdash^{l}_{\mathbf{V}\mathfrak{C}} F(\varphi, \varphi)$. Furthermore, observe that, by assumption, we have $\models_{\mathbf{V}\mathfrak{C}} \varphi = \psi$, hence it must be the case that

$$\models_{\mathbf{V}\mathfrak{C}} \delta(\varphi, \varphi) = \delta(\varphi, \psi)$$

Where $\delta(\varphi, \psi)$ is a formula in the language $\mathcal{L}_{\Box\!\!\rightarrow}$ obtained by combining φ and ψ, and the same holds for $\delta(\varphi, \varphi)$. Hence, in particular, it is the case that

$$\models_{\mathbf{V}\mathfrak{C}} F(\varphi, \varphi) = F(\varphi, \psi)$$

Then, by Corollary 1.1, it follows that

$$F(\varphi, \varphi) \dashv\vdash^{l}_{\mathbf{V}\mathfrak{C}} F(\varphi, \psi)$$

and since $\vdash^{l}_{\mathbf{V}\mathfrak{C}} F(\varphi, \varphi)$, it must be the case that

$$\vdash^{l}_{\mathbf{V}\mathfrak{C}} F(\varphi, \psi)$$

Since $\vdash^{l}_{\mathbf{V}\mathfrak{C}}$ is algebraizable by assumption, we have that $\models_K E(F(\varphi, \psi))$ and so

$$\models_K \varphi = \psi$$

Hence, every equation that holds in $\mathbf{V}\mathfrak{C}$ will also hold in K. Therefore K must be a class of $\mathbf{V}\mathfrak{C}$ conditional algebras.

□

Now, we recap a general theorem concerning algebraizability:

Notation 1.5.

Given an algebra $\mathbf{A}$, *and a logic* $\vdash$, *let* K *be a class of algebras of the same type of* $\mathbf{A}$, *i.e. in the same signature as* $\mathbf{A}$

- $Fi_{\vdash}(\mathbf{A})$ *denotes the lattice of deductive filters of* $\vdash$ *on* $\mathbf{A}$, *ordered by set inclusion*
- $Con_K(\mathbf{A})$ *denotes the lattice of* K*-relative congruences on* $\mathbf{A}$, *ordered by set inclusion.*

  A congruence θ *on* $\mathbf{A}$ *is* K*-relative if* $A_{/\theta} \in K$, *i.e. the quotient of* $\mathbf{A}$ *modulo* θ *is in* K.
- $Fi_{\vdash}(\mathbf{A}) \rightarrowtail\!\!\!\!\twoheadrightarrow Con_K(\mathbf{A})$ *means that* $Fi_{\vdash}(\mathbf{A})$ *and* $Con_K(\mathbf{A})$ *are in bijective correspondence*

Theorem 1.7 (Blok and Pigozzi 2014)**.** *Point 1 implies point 2 below:*

1. *A logic* $\vdash$ *is algebraizable with respect to a class of algebras* K
2. *for every algebra* $\mathbf{A}$ *in the same signature as the algebras in* K, $Fi_{\vdash}(A) \rightarrowtail\!\!\!\!\twoheadrightarrow Con_K(A)$, *i.e. the lattice of deductive filter of* $\vdash$ *on* $\mathbf{A}$ *is in bijective correspondence with the lattice of congruences on* $\mathbf{A}$

From the above observations, the following result follows:

Theorem 1.8. *Let* $\mathfrak{C}$ *be a condition or a family of conditions among those in* $\{$**(N)**, **(T)**, **(W)**, **(C)**, **(S)**, **(U)**, **(U-)**, **(UT)**$\}$; *then, the local* $\mathbf{V}\mathfrak{C}$ *logic is not algebraizable*

Proof. For reductio, assume $\vdash^l_{\mathbf{V}\mathfrak{C}}$ is algebraizable. Then, by Lemma 1.7, $\vdash^l_{\mathbf{V}\mathfrak{C}}$ would be algebraizable with respect to a class of conditional algebras K satisfying condition $\mathfrak{C}$. Now, observe that the algebra $\mathbf{V}$ in Lemma 1.6 is a conditional algebra, hence it is of the same type as those the class K. Therefore, by Theorem 1.7, it must follows that

$$Fi_{\vdash^l_{\mathbf{V}\mathfrak{C}}}(\mathbf{V}) \rightarrowtail\!\!\!\!\twoheadrightarrow Con_K(\mathbf{V})$$

However, notice that $Fi_{\vdash^l_{\mathbf{V}\mathfrak{C}}}(\mathbf{V})$ contains four elements:

- $\{W\}$
- $\{\{w_1\}, W\}$

- $\{\{w_2\}, W\}$
- $\{\{w_1\}, \emptyset, \{w_2\}, W\}$

Indeed, it is easy to show that each of the set above is a deductive filter of $\vdash^{l}_{\mathbf{V}\mathfrak{C}}$ on $\mathbf{V}$ since $\mathbf{V}$ satisfies all conditions $\mathfrak{C}$. However, $\mathbf{V}$ has only two congruences:

- the trivial one $\{(a, b) \mid a, b \in V\}$, where all the elements are congruent to each other
- the minimal one, $\{(a, a) \mid a \in V\}$, where each element is only congruent to itself

The above observation follows from the following reasoning. First, observe that for any congruence θ on $\mathbf{V}$, if $(W, \emptyset) \in \theta$, then θ is the trivial congruence. Consider the minimal congruence θ on $\mathbf{V}$. If $(\{w_1\}, W) \in \theta$, then, since θ must preserve $\boxdot$, we have that $(\emptyset, W) \in \theta$ too, and this follows from Lemma 1.6, since $\boxdot\{w_1\} = \emptyset$ and $\boxdot W = W$. Therefore, θ would be the trivial congruence. We can reason analogously for all the other congruences, ending up with the observation that $Con_K(\mathbf{V})$ contains only the trivial and the minimal congruence. Hence,

$$Fi_{\vdash^{l}_{\mathbf{V}\mathfrak{C}}}(\mathbf{V}) \not\twoheadrightarrow Con_K(\mathbf{V})$$

And this is in contradiction with out initial observation. Therefore $\vdash^{l}_{\mathbf{V}\mathfrak{C}}$ is not algebraizable.

□

From the above results, it follows that not all local variably strict conditional logic are algebraizable. In what follows, we will analyze the algebrizability of the remaining local logics **(VWA)**, **(VTSA)**, **(VTA)**, **(VNSA)**, **(VNA)**, **(VSA)**, **(VA)**. For these logics, we can reason similarly to the case illustrated above. The strategy is the same: we show that there is a conditional algebra in which the lattice of deductive filter is not in bijective correspondence with the lattice of relative congruences over that algebra. In particular, the key lemma would be the following:

Lemma 1.8. *Consider a finite spherical Lewisian model $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$ where:*

- $W = \{w_1, w_2\}$
- *$\mathcal{S}$ is such that:*
    - $\mathcal{S}(w_1) = \mathcal{S}(w_2) = \{\{w_1\}, \{w_1, w_2\}\}$
- *$\vDash$ is such that:*

- $w_1 \vDash p$ *and* $w_2 \nvDash p$
- $w_1 \nvDash q$ *and* $w_2 \nvDash q$

Then the structure $\langle \wp(W), \cap, \cup, ^-, \Box\!\!\rightarrow, W, \emptyset \rangle$ *is as conditional algebra, in which*

- $\cap$ *is set-theoretic intersection*
- $\cup$ *is set-theoretic union*
- $^-$ *is set-theoretic complement*
- $\Box\!\!\rightarrow : \wp(W) \times \wp(W) \rightarrow \wp(X)$ *is defined as follows: for all* $X, Y \in \wp(W)$,

$$X \Box\!\!\rightarrow Y = \{w \in W \mid min^X_{\subseteq}(\mathcal{S}(w)) \subseteq Y\}$$

where

$$min^X_{\subseteq}(\mathcal{S}(w)) = \begin{cases} \text{the minimal } S \in \mathcal{S}(w) \text{ such that } S \cap X \neq \emptyset & \text{if } \bigcup \mathcal{S}(w) \cap X \neq \emptyset \\ \emptyset & \text{otherwise} \end{cases}$$

and, moreover,

- $\boxdot\{w_1\} = \boxdot\{w_2\} = \boxdot\emptyset = \emptyset$,
- $\boxdot W = W$

The spherical Lewisian model above satisfies constraints **(T)**, **(S)**, **(N)**, **(A)**, *as well as its corresponding algebra.*

Proof. The proof proceeds analogously to the proof of Lemma 1.6. □

The above example can be employed to show that the local logics **(VWA)**, **(VTA)**, **(VNA)**, **(VA)** are not algebraizable, analogously to Theorem 1.8. Additionally, the following lemma holds:

Lemma 1.9. *Consider a finite spherical Lewisian model* $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$ *where:*

- $W = \{w_1, w_2\}$
- $\mathcal{S}$ *is such that:*
  - $\mathcal{S}(w_1) = \mathcal{S}(w_2) = \{W\}$
- $\vDash$ *is such that:*
  - $w_1 \vDash p$ *and* $w_2 \nvDash p$
  - $w_1 \nvDash q$ *and* $w_2 \nvDash q$

Then the structure $\langle \wp(W), \cap, \cup, ^{-}, \Box\!\!\rightarrow, W, \emptyset \rangle$ *is a conditional algebra, in which*

- $\cap$ *is set-theoretic intersection*
- $\cup$ *is set-theoretic union*
- $^{-}$ *is set-theoretic complement*
- $\Box\!\!\rightarrow : \wp(W) \times \wp(W) \to \wp(X)$ *is defined as follows: for all* $X, Y \in \wp(W)$,

$$X \Box\!\!\rightarrow Y = \{w \in W \mid min^X_{\subseteq}(\mathcal{S}(w)) \subseteq Y\}$$

where

$$min^X_{\subseteq}(\mathcal{S}(w)) = \begin{cases} \text{the minimal } S \in \mathcal{S}(w) \text{ such that } S \cap X \neq \emptyset & \text{if } \bigcup \mathcal{S}(w) \cap X \neq \emptyset \\ \emptyset & \text{otherwise} \end{cases}$$

and, moreover,

- $\boxdot\{w_1\} = \boxdot\{w_2\} = \boxdot\emptyset = \emptyset$,
- $\boxdot W = W$

The above spherical Lewisian model satisfies constraint **(W)**, **(T)**, **(N)**, **(A)**, *as well as its corresponding algebra.*

The above lemma can be used to show that the local logics **(VWA)**, **(VTA)**, **(VNA)**, **(VA)** are not algebraizable, analogously to Theorem 1.8. The only remaining case is the local logic **WCA**. However, it is not difficult to show that the global **WCA** logic coincides with the local **WCA**.

Theorem 1.9. *For* $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}_{\Box\!\rightarrow}}$,

$$\Gamma \models^l_{\mathbf{VCA}} \varphi \Leftrightarrow \Gamma \models^g_{\mathbf{VCA}} \varphi$$

Proof.

($\Rightarrow$) Straightforward by Remark 1.2

($\Rightarrow$) By contraposition, assume $\Gamma \not\models^g_{\mathbf{VCA}} \varphi$, then there is **WCA** spherical Lewisian model $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$ such that $\mathcal{M} \vDash \gamma$ for all $\gamma \in \Gamma$ but $\mathcal{M} \nvDash \varphi$, namely there is a $w \in W$ such that $w \nvDash \varphi$. Now, consider the sub-model $\mathcal{M}'$ generated by $\{w\}$. By Lemma 1.2, $\mathcal{M}'$ only contains w and preserves truth of formulas. Hence, it would be the case that $\mathcal{M}' \vDash \gamma$ for all $\gamma \in \Gamma$ but $\mathcal{M}' \nvDash \varphi$. This implies that $\Gamma \not\models^l_{\mathbf{VCA}} \varphi$

□

Therefore, by Corollary 1.2, $\vdash^l_{\mathbf{VCA}}$ is algebraizable. In conclusion, all the local variably strict conditional logics are not algebraizable, except for **WCA**.

1.3 Conclusions

In this chapter we have distinguished between a local and global counterpart of variably strict conditional logics and proved some interesting connections between the two in a way that mirrors the connection between global and local modal logics. These two notions - global and local - lay the foundation for a systematic algebraic treatment of variably strict conditional logics.

To begin with, we introduced a class of algebras known as conditional algebras, where global variably strict conditional logics coincide with equational consequences, and local variably strict conditional logics correspond to the preservation of degrees of truth. We also provided proofs regarding the algebraizability of these logics. Specifically, we demonstrated that global variably strict conditional logics can be algebraized with respect to a corresponding class of conditional algebras. On the other hand, with the sole exception of the logic **WCA**, local variably strict conditional logics cannot be algebraized.

It's essential to note that this chapter covered fundamental results that lay the groundwork for an extensive and systematic treatment of variably strict conditional logics. The next crucial step in this direction would involve delving into the structure theory of conditional algebras. Particularly, the characterization of deductive filters for global variably strict conditional logics remains an open problem. Additionally, developing a duality theory for conditional algebras would enable us to establish connections between conditional algebras and Lewisian possible worlds models for variably strict conditionals. This avenue of research would significantly contribute to our understanding of variably strict conditional logics and their underlying semantic structures in a broader context.

Chapter 2

Boolean Algebras of Counterfactuals

In this chapter, we present an algebraic approach to Lewisian counterfactuals, using the innovative framework of Boolean algebras of conditionals, recently introduced by Flaminio, Godo, and Hosni (2020). We begin by reviewing the BACs framework and discussing its philosophical implications on the debate over conditionals and their probabilities. Subsequently, we explore modal extensions of the BACs, within which Lewisian counterfactuals can be precisely defined. Finally, we conclude with some reflections on the philosophical significance of the results we have obtained.

2.1 Background

We need some additional technical background extending the notions introduced in Chapter 1. Specifically, we will define a new class of Lewisian models and use a new object language.

2.1.1 Total Models

First, we need some notation:

- for $\varphi \in For_{\mathcal{L}}$, i.e. for φ being a classical formula, $\vdash_{CPL} \varphi$ means that φ is a theorem of classical logic, or, equivalently, $\models_{CPL} \varphi$ means that φ is a classical tautology

We are ready to introduce a new class of Lewisian models:

Definition 2.1. *A total spherical Lewisian model* $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$ *is a spherical Lewisian model satisfying the following condition:*

- ***Totality***: *for all* $\varphi \in For_{\mathcal{L}}$*, for all* $w \in W$*, if* $\nvdash_{CPL} \neg\varphi$*, then there is a* $v \in \bigcup \mathcal{S}(w)$ *such that* $v \vDash \varphi$ *(or equivalently* $[\varphi] \cap \bigcup \mathcal{S}(w) \neq \emptyset$*).*

Namely, a total spherical Lewisian model is a spherical Lewisian model where all satisfiable classical formulas are made true by at least one world in each system of spheres. This condition ensures that counterfactuals with classical satisfiable antecedents are not vacuously true in a world. The totality constraint is characterized by an axiom schema establishing that each would-counterfactual having a classical satisfiable antecedent implies its corresponding might-counterfactual, as the following result shows.

Lemma 2.1. *A spherical Lewisian model* $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$ *satisfies the totality constraint if and only if for all* $w \in W$*, for all* $\varphi \in For_{\mathcal{L}}$ *such that* $\nvdash_{CPL} \neg\varphi$*,* $w \vDash (\varphi \mathrel{\Box\!\!\rightarrow} \psi) \supset (\varphi \mathrel{\Diamond\!\!\rightarrow} \psi)$*.*

Proof.

($\Rightarrow$) Consider a total spherical Lewisian model $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$ and a formula $\varphi \in For_{\mathcal{L}}$ such that $\nvdash_{CPL} \neg\varphi$. Take an arbitrary $w \in W$; by the totality condition we know that $[\varphi] \cap \bigcup \mathcal{S}(w) \neq \emptyset$. Assume $w \vDash \varphi \mathrel{\Box\!\!\rightarrow} \psi$, then, by totality, there is a $S \in \mathcal{S}(w)$ such that $[\varphi] \cap S \subseteq [\psi]$. This implies that for all $S \in \mathcal{S}(w)$, if $[\varphi] \cap S \neq \emptyset$ then $[\varphi] \cap [\psi] \cap S \neq \emptyset$, hence $w \vDash \varphi \mathrel{\Diamond\!\!\rightarrow} \psi$.

($\Leftarrow$) For the other direction, let's reason for reductio and assume that $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$ is a total, but for some $w \in W$, there is a $\varphi \in For_{\mathcal{L}}$ such that $\nvdash_{CPL} \neg\varphi$ and $w \nvDash (\varphi \mathrel{\Box\!\!\rightarrow} \psi) \supset (\varphi \mathrel{\Diamond\!\!\rightarrow} \psi)$. Then $w \vDash \varphi \mathrel{\Box\!\!\rightarrow} \psi$ and $w \nvDash \varphi \mathrel{\Diamond\!\!\rightarrow} \psi$. By the totality condition, we know that $[\varphi] \cap \bigcup \mathcal{S}(w) \neq \emptyset$. Thus, by semantic conditions, we have that there are $S, S' \in \mathcal{S}(w)$ such that $[\varphi] \cap S \subseteq [\psi]$ and $[\varphi] \cap S' \nsubseteq [\psi]$. By the nestedness condition of spherical Lewisian models, we know that either $S \subseteq S'$ or $S' \subseteq S$, in both cases we reach a contradiction.

□

Analogously, we can introduce the corresponding class of functional Lewisian models:

Definition 2.2. *A total functional Lewisian model* $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$ *is a functional Lewisian model satisfying the following condition:*

- ***Totality***: *for all* $\varphi \in For_{\mathcal{L}}$*, for all* $w \in W$*, if* $\nvdash_{CPL} \neg\varphi$*, then* $f(\varphi, w) \neq \emptyset$

The same axiom schema as before characterizes the class of total functional Lewisian models:

Lemma 2.2. *A functional Lewisian model* $\mathcal{M} = \langle W, f, \vDash \rangle$ *satisfies the totality constraint if and only if for all* $w \in W$*, for all* $\varphi \in For_{\mathcal{L}}$ *such that* $\nvdash_{CPL} \neg\varphi$*,* $w \vDash (\varphi \,\Box\!\!\rightarrow \psi) \supset (\varphi \,\Diamond\!\!\rightarrow \psi)$

Proof.

($\Rightarrow$) Assume $\mathcal{M} = \langle W, f, \vDash \rangle$ satisfies the totality constraint and assume $w \vDash \varphi \,\Box\!\!\rightarrow \psi$ for an arbitrary $w \in W$ and an arbitrary $\varphi \in For_{\mathcal{L}}$ such that $\nvdash_{CPL} \neg\varphi$. By semantic conditions, we have that $f(\varphi, w) \subseteq [\psi]$ and, since $f(\varphi, w) \neq \emptyset$, by totality, we have that $f(\varphi, w) \cap [\psi] \neq \emptyset$, hence $w \vDash \varphi \,\Diamond\!\!\rightarrow \psi$. Therefore, $w \vDash (\varphi \,\Box\!\!\rightarrow \psi) \supset (\varphi \,\Diamond\!\!\rightarrow \psi)$.

($\Leftarrow$) By contraposition, assume that $\mathcal{M} = \langle W, f, \vDash \rangle$ doesn't satisfy the totality condition; thus for some $\varphi \in For_{\mathcal{L}}$ such that $\nvdash_{CPL} \neg\varphi$, for some $w \in W$, $f(\varphi, w) = \emptyset$. Hence, by semantic conditions, $w \vDash \varphi \,\Box\!\!\rightarrow \psi$ but $w \nvDash \varphi \,\Diamond\!\!\rightarrow \psi$, namely $w \nvDash (\varphi \,\Box\!\!\rightarrow \psi) \supset (\varphi \,\Diamond\!\!\rightarrow \psi)$.

□

The following table summarizes the features of total models:

	Spherical Lewisian Models	
	Condition	Axiom
($^+$) Totality	for all $\varphi \in For_{\mathcal{L}}$ such that $\nvdash_{CPL} \neg\varphi$, $[\varphi] \cap \bigcup \mathcal{S}(w) \neq \emptyset$	for all $\varphi \in For_{\mathcal{L}}$ such that $\nvdash_{CPL} \neg\varphi$, $(\varphi \,\Box\!\!\rightarrow \psi) \supset (\varphi \,\Diamond\!\!\rightarrow \psi)$

	Functional Lewisian Models	
	Condition	Axiom
($^+$) Totality	for all $\varphi \in For_{\mathcal{L}}$ such that $\nvdash_{CPL} \neg\varphi$, $f(\varphi, w) \neq \emptyset$	for all $\varphi \in For_{\mathcal{L}}$ such that $\nvdash_{CPL} \neg\varphi$, $(\varphi \,\Box\!\!\rightarrow \psi) \supset (\varphi \,\Diamond\!\!\rightarrow \psi)$

Table 2.1: The table schematically summarizes the totality condition over functional and spherical Lewisian models and its characteristic axiom.

Notation 2.1.

- *Consider any variably strict conditional logic* **V**$\mathfrak{C}$ *among those in Definition 0.4. Then* **V**$\mathfrak{C}^+$ *is the logic resulting by adding axiom* $^+$ *in Table 2.1.1 to* **V**$\mathfrak{C}$*. For instance* **VC**$^+$ *is the logic resulting by adding* $^+$ *to* **VC**.

Theorem 0.1 can be extended in order to include the totality axioms:

Theorem 2.1. *Let $\mathfrak{C}$ be an axiom/condition or a family of axioms/conditions (possibly empty) among those in Tables 1 and 2.1.1, (i.e. $\{\mathbf{N}, \mathbf{T}, \mathbf{W}, \mathbf{C}, \mathbf{S}, \mathbf{A}, \mathbf{U},^{+}\}$), then, for all $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}_{\Box\!\rightarrow}}$, the following holds:*

$$\Gamma \vdash_{\mathbf{V}\mathfrak{C}} \varphi \Leftrightarrow \Gamma \models_{\mathbf{V}\mathfrak{C}} \varphi \Leftrightarrow \Gamma \models_{\mathbf{V}^{\mathfrak{f}}\mathfrak{C}} \varphi$$

Proof. The result follows easily from an adaptation of the soundness and completeness proofs in (Lewis 1971, 1973b) combined with Lemma 2.2. □

2.1.2 New Languages

In this chapter, we need to present the language of the *LBC*-logic (Logic of Boolean Conditionals), denoted as $\mathcal{L}_{LBC}$, introduced by Flaminio, Godo, and Hosni (2020). Let us assume, for the length of this chapter, that our underlying set of propositional variables *Var* is finite, i.e. $Var = \{p_1, p_2, \ldots, p_n\}$. Thus, the languages we introduced in Chapter 1, i.e. $\mathcal{L}$ and $\mathcal{L}_{\Box\!\rightarrow}$, are now meant to be built upon a finite *Var*. Upper-case Greek letters indicate formulas in the language $\mathcal{L}_{LBC}$.

Definition 2.3 (*LBC* language)**.** *$\mathcal{L}_{LBC}$ is a language obtained by expanding $\mathcal{L}$ with the binary connective $(\cdot \mid \cdot)$, understood as a probabilistic conditional such that $(\varphi \mid \psi)$ can be read as "φ given ψ". Formulas of $\mathcal{L}_{LBC}$ are defined as follows:*

- *if φ, ψ are formulas of $\mathcal{L}$ and $\nvdash_{CPL} \neg\varphi$, then $(\psi \mid \varphi)$ is a formula of $\mathcal{L}_{LBC}$;*
- *if Φ, Ψ are formulas of $\mathcal{L}_{LBC}$, then $\neg\Phi$, $\Phi \wedge \Psi$ and $\Phi \vee \Psi$ are formulas of $\mathcal{L}_{LBC}$;*
- *nothing else is a formula of $\mathcal{L}_{LBC}$.*

$For_{\mathcal{L}_{LBC}}$ denotes the set of formulas of $\mathcal{L}_{LBC}$.

Namely, formulas in *LBC* are all conditional formulas of the form $(\varphi \mid \psi)$ where φ and ψ are classical formulas from $\mathcal{L}$. Moreover, it will be useful to isolate a particular fragment of Lewis' language $\mathcal{L}_{\Box\!\rightarrow}$ which is structurally very similar to $\mathcal{L}_{LBC}$:

Definition 2.4 ($\mathcal{L}^{\restriction}_{\Box\!\rightarrow}$ Language)**.** *$\mathcal{L}^{\restriction}_{\Box\!\rightarrow}$ is a language obtained by restricting $\mathcal{L}_{\Box\!\rightarrow}$ to its first-degree formulas, i.e. formulas not containing embedded conditionals. Specifically, formulas of $\mathcal{L}^{\restriction}_{\Box\!\rightarrow}$ are defined as follows:*

- *if φ is a formula of $\mathcal{L}$, then φ is a formula of $\mathcal{L}^{\restriction}_{\Box\!\rightarrow}$;*
- *if φ, ψ are formulas of $\mathcal{L}$ and $\nvdash_{CPL} \neg\varphi$, then $\varphi \mathrel{\Box\!\!\rightarrow} \psi$ is a formula of $\mathcal{L}^{\restriction}_{\Box\!\rightarrow}$;*
- *if φ, ψ are formulas of $\mathcal{L}^{\restriction}_{\Box\!\rightarrow}$, then so are $\neg\varphi, \varphi \wedge \psi$, and $\varphi \vee \psi$*
- *nothing else is a formula of $\mathcal{L}^{\restriction}_{\Box\!\rightarrow}$.*

Namely, formulas of $\mathcal{L}^{\upharpoonright}_{\Box\!\!\rightarrow}$ include classical formulas and those conditional formulas in $\mathcal{L}_{\Box\!\!\rightarrow}$ having a satisfiable classical antecedent and not containing any nested conditionals. Though obvious, we briefly recap the definition of classical valuations, since we need those for interpreting $\mathcal{L}_{LBC}$.

Definition 2.5. *A classical valuation v if a function $v : Var \rightarrow \{0, 1\}$ from our (finite) set of propositional variables to the set of Boolean values 0 and 1. v is extended to compound formulas of $\mathcal{L}$ as follows:*

$$\begin{array}{lcl} v(\neg\varphi) = 1 & \Leftrightarrow & v(\varphi) = 0 \\ v(\varphi \wedge \psi) = 1 & \Leftrightarrow & v(\varphi) = 1 \textit{ and } v(\psi) = 1 \\ v(\varphi \wedge \psi) = 1 & \Leftrightarrow & v(\varphi) = 1 \textit{ or } v(\psi) = 1 \end{array}$$

Furthermore, Val_{CPL} denotes the set of all classical valuations of our language $\mathcal{L}$.

Additionally, we use the following notation:

Notation 2.2.

- *For $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}}$,*

$$\begin{array}{lcl} \Gamma \models_{CPL} \varphi & \Leftrightarrow & \textit{for all classical valuations } v, \\ & & \textit{if } v(\gamma) = 1 \textit{ for all } \gamma \in \Gamma, \textit{ then } v(\varphi) = 1 \end{array}$$

It is well known that classical logic (*CPL*) is sound and complete with respect to classical valuations:

Theorem 2.2. *For all $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}}$, $\Gamma \models_{CPL} \varphi \Leftrightarrow \Gamma \vdash_{CPL} \varphi$.*

2.1.3 Restricted Models

In this section, we conduct a technical investigation of the properties of spherical Lewisian models. The results we demonstrate in this section have exclusively a technical purpose aiming at showing soundness and completeness of the logic $\mathbf{VC}^+$ over the restricted language $\mathcal{L}^{\upharpoonright}_{\Box\!\!\rightarrow}$ with respect to spherical Lewisian models satisfying totality. Recall that in this chapter we are assuming a finite number of propositional variables, $Var = \{p_1, \ldots, p_n\}$. Moreover, let us establish some notation:

Notation 2.3.

- *For a Kripke frame* $\langle W, R \rangle$ *for any* $w \in W$, $R[w] = \{v \mid v \in W$ *and* $wRv\}$ *denotes the set of all the accessible worlds from* w*;*
- *Given a set* X *and an equivalence relation* $\equiv$, $X_{/\equiv} = \{[x]_\equiv \mid x \in X\}$ *is the quotient set obtained from* X *and* $\equiv$*, namely* $X_{/\equiv}$ *is the set of* $\equiv$*-equivalence classes over* X*; and* $[x]_\equiv$ *denotes the* $\equiv$*-equivalence class of the element* x*.*
- *for a set* X, $|X|$ *denotes the cardinality of* X

Definition 2.6. *For a total spherical Lewisian model* $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$*, the restricted spherical model associated to* $\mathcal{M}$*, is a tuple* $\mathcal{M}_\approx = \langle W_\approx, \mathcal{S}_\approx, \vDash_\approx, c \rangle$ *where:*

- $\approx \subseteq W \times W$ *is a binary relation over* W *defined as follows:*

$$w \approx v \Leftrightarrow \text{ for all } \varphi \in For_{\mathcal{L}}, (w \vDash \varphi \Leftrightarrow v \vDash \varphi)$$

  that is, w *and* v *are* $\approx$*-equivalent iff they force exactly the same classical formulas. Clearly* $\approx$ *is an equivalence relation;*

- $W_\approx = \{[w]_\approx \mid w \in W\}$ *is the set of* $\approx$*-equivalence classes of* W*;*
- $c : W_\approx \to W$ *is a choice function that select a representative element* $w \in W$ *for each equivalence class in* $W_\approx$*, e.g. for* $w \in W$, $c([w]_\approx) \in [w]_\approx$
- $\mathcal{S}_\approx : W_\approx \to \wp(\wp(W_\approx)) \setminus \emptyset$ *is defined as follows:*

$$\text{for all } [w]_\approx \in W_\approx, \mathcal{S}_\approx([w]_\approx) = \{S_\approx \mid S \in \mathcal{S}(c([w])_\approx)\}$$

- $\vDash_\equiv \subseteq W_\approx \times Var$ *is a valuation relation defined as follows:*

$$[w]_\approx \vDash_\approx p \Leftrightarrow c([w]_\approx) \vDash_\approx p$$

Moreover we set $[\varphi]_\approx = \{[w]_\approx \mid [w]_\approx \vDash_\approx \varphi\}$

The choice function in a restricted model $\mathcal{M}_\equiv$ is needed in order for $\mathcal{S}_\approx$ and $\vDash_\approx$ to be well defined. In fact, the system of spheres associated with an equivalence class $[w]_\approx$ will be built out of the systems of spheres in the original model $\mathcal{M}$. However, in $[w]_\approx$ there are many several elements of W, hence the systems of spheres associated to $[w]_\approx$ will be obtained from the original system of spheres associated to a chosen one (by the function c) among the elements of $[w]_\approx$.

Some properties of the restricted models will prove useful:

Lemma 2.3. *For a total spherical Lewisian model* $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$*, consider its associated restricted spherical model* $\mathcal{M}_{\approx} = \langle W_{\approx}, \mathcal{S}_{\approx}, \vDash_{\approx}, c \rangle$*. The following holds:*

1. $\mathcal{M}_{\approx}$ *is a spherical total Lewisian model;*
2. *if* $\mathcal{M}$ *satisfies* ***Centering****, then* $\mathcal{M}_{\approx}$ *satisfies* ***Centering*** *too;*
3. *assume* $|Var| = n$ *for some* $n \in \mathbb{N}$*, then for all* $w \in W$*,* $|\bigcup \mathcal{S}_{\approx}([w]_{\approx})| = |W_{\approx}| = 2^n = |Val_{CPL}|$*; namely the cardinality of* $W_{\approx}$ *amounts to the number of classical valuations of* $\mathcal{L}$*. Hence, there is a bijection* $E : Val_{CPL} \rightarrowtail\!\!\!\!\twoheadrightarrow W_{\approx}$ *from elements of* $W_{\approx}$ *and classical valuations defined as:*

$$E(v) = \{w \in W \mid w \vDash p \Leftrightarrow v(p) = 1 \textit{ and } w \nvDash p \Leftrightarrow v(p) = 0\}$$

Proof. The proof is rather tedious and included in the Appendix B □

The following example will clarify the construction involved in the restricted spherical model and provide some graphical intuitions.

Example of Restricted Model

Assume $Var = \{p, q\}$ and let $\langle W, \mathcal{S}, \vDash \rangle$ be a total sphere model with:

- $W = \{w_1, w_2, w_3, w_4, w_5, w_6\}$
- $\mathcal{S}(w_1) = \{\{w_1\}, \{w_1, w_3, w_4, w_5\}, \{w_1, w_2, w_3, w_4, w_5, w_6\}\}$
- $[p \wedge q] = \{w_1, w_3\}$, $[\neg p \wedge q] = \{w_4, w_5\}$, $[p \wedge \neg q] = \{w_2\}$, $[\neg p \wedge \neg q] = \{w_6\}$

Consider now $\langle W_{\approx}, \mathcal{S}_{\approx}, \vDash_{\approx} \rangle$, with $w_1 \approx w_3$ and $w_4 \approx w_5$ and the underlying choice function c such that $c([w_1]_{\approx}) = w_1$. Thus:

- $W_{\approx} = \{[w_1]_{\approx}, [w_2]_{\approx}, [w_4]_{\approx}, [w_6]_{\approx}\}$
- $\mathcal{S}_{\approx}([w_1]_{\approx}) = \mathcal{S}(c[w_1]_{\approx})_{/\approx} = \mathcal{S}(w_1)_{/\approx} = \{\{[w_1]_{\approx}\}, \{[w_1]_{\approx}, [w_4]_{\approx}\}, \{[w_1]_{\approx}, [w_2]_{\approx}, [w_4]_{\approx}, [w_6]_{\approx}\}\}$
- $[p \wedge q]_{\approx} = \{[w_1]_{\approx}\}$, $[\neg p \wedge q]_{\approx} = \{[w_4]_{\approx}\}$, $[p \wedge \neg q]_{\approx} = \{[w_2]_{\approx}\}$, $[\neg p \wedge \neg q]_{\approx} = \{[w_6]_{\approx}\}$

This example is depicted in the following figure:

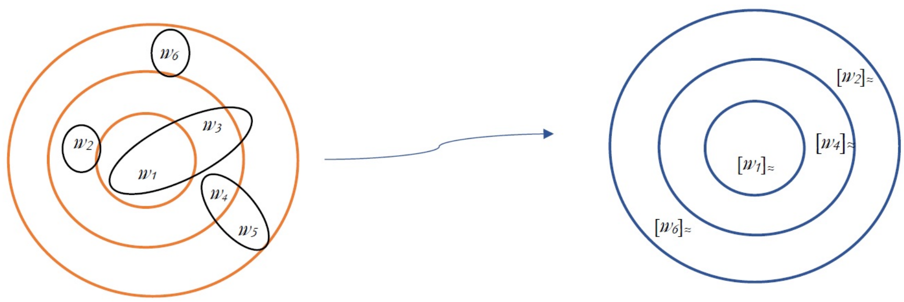


Figure 2.1: The figure depicts how a system of spheres would change when passing from a total sphere model (on the left) to its corresponding restricted model (on the right). The black circles select the worlds in the same equivalence class.

Truth of formulas in $\mathcal{L}^{\restriction}_{\Box\!\!\rightarrow}$ is preserved when passing from a total spherical Lewisian model to a restricted spherical model:

Lemma 2.4. *For a total sphere model $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$, consider its associated restricted spherical model $\mathcal{M}_{\approx} = \langle W_{\equiv}, \mathcal{S}_{\approx}, \vDash_{\approx}, c \rangle$. For all formulas $\varphi \in For_{\mathcal{L}^{\restriction}_{\Box\!\!\rightarrow}}$, the following holds:*

$$\mathcal{M}, c([w]_{\approx}) \vDash \varphi \Leftrightarrow \mathcal{M}_{\approx}, [w]_{\approx} \vDash_{\approx} \varphi$$

Proof. By induction. We show one case for exemplification. Assume $c([w]_{\approx}) = w$. Consider $\varphi = \delta \Box\!\!\rightarrow \psi$. Assume $\mathcal{M}, w \vDash \delta \Box\!\!\rightarrow \psi$. By totality and by semantic conditions, there is a $S \in \mathcal{S}(w)$ and a $v \in S$ such that $v \vDash \delta$ and for all $x \in S$, $x \vDash \delta \supset \psi$. By definition of $\mathcal{M}_{\approx}$, and by induction hypothesis, we have that $S_{\approx} \in \mathcal{S}_{\approx}([w]_{\approx})$, $[v]_{\approx} \in S_{\approx}$ and $[v]_{\approx} \vDash_{\approx} \delta$ and moreover for all $[x]_{\approx} \in S_{\approx}$, $[x]_{\approx} \vDash_{\approx} \delta \supset \psi$. Thus $[w]_{\approx} \vDash_{\approx} \delta \Box\!\!\rightarrow \psi$. Similarly for the other direction. □

Finally, one last technical tool will be essential to simplify our technical results:

Definition 2.7. *Given a restricted spherical model $\mathcal{M}_{\approx} = \langle W_{\approx}, \mathcal{S}_{\approx}, \vDash_{\approx}, c \rangle$, its corresponding restricted canonical model is a tuple $\mathcal{M}^{\mathfrak{C}} = \langle Val_{CPL}, \mathcal{S}^{\mathfrak{C}}, \vDash^{\mathfrak{C}} \rangle$ where:*

- *$\mathcal{S}^{\mathfrak{C}} : Val_{CPL} \to \wp(\wp(Val_{CPL}))$ is defined as follows: for all $S \in \mathcal{S}_{\approx}(E(v))$, define $S^{\mathfrak{C}} = \{E^{-1}(x) \mid x \in S\}$. Namely, $S^{\mathfrak{C}}$ is the set of classical valuations naturally associated by E to the elements of S. Then, for all $v \in Val_{CPl}$,*

$$\mathcal{S}^{\mathfrak{C}}(v) = \{S^{\mathfrak{C}} \mid S \in \mathcal{S}_{\approx}(E(v))\}$$

- *$\vDash^{\mathfrak{C}} \subseteq Val_{CPL} \times Var$ is defined as follows: for all $v \in Val_{CPL}$, for all $p \in Var$,*

$$v \vDash^{\mathfrak{C}} p \Leftrightarrow v(p) = 1$$

and $\vDash^{\mathfrak{C}}$ *is extended to all formulas of* $\mathcal{L}_{\Box\!\rightarrow}$ *according to Definition* 0.2*. Furthermore, we set:* $[\varphi]^{\mathfrak{C}} = \{v \in Val_{CPL} \mid v \vDash \varphi\}$

The following lemma is readily provable by construction, by Lemma 2.3, and by induction:

Lemma 2.5. *For a restricted spherical model* $\mathcal{M}_{\approx} = \langle W_{\approx}, \mathcal{S}_{\approx}, \vDash_{\approx}, c\rangle$*, consider its corresponding restricted canonical model* $\mathcal{M}^{\mathfrak{C}} = \langle Val_{CPL}, \mathcal{S}^{\mathfrak{C}}, \vDash^{\mathfrak{C}}\rangle$*. Then, the following holds:*

- $\mathcal{M}^{\mathfrak{C}}$ *is indeed a total spherical Lewisian model;*
- *for all* $v \in Val_{CPL}$*,* $\bigcup \mathcal{S}^{\mathfrak{C}}(v) = Val_{CPL}$*;*
- *if* $\mathcal{M}_{\approx}$ *satisfies* ***Centering****, then* $\mathcal{M}^{\mathfrak{C}}$ *satisfies* ***Centering*** *too;*
- *for all* $\varphi \in For_{\mathcal{L}^{\restriction}_{\Box\!\rightarrow}}$*, for all* $x \in W_{\approx}$*,*

$$\mathcal{M}_{\approx}, x \vDash_{\approx} \varphi \Leftrightarrow \mathcal{M}^{\mathfrak{C}}, E(x) \vDash^{\mathfrak{C}} \varphi$$

From Lemma 2.5 and Lemma 2.4, the next result readily follows:

Corollary 2.1. *Consider a total spherical model* $\mathcal{M} = \langle W, \mathcal{S}, \vDash\rangle$*, for all* $\varphi \in For_{\mathcal{L}^{\restriction}_{\Box\!\rightarrow}}$*, for all* $x \in W_{\approx}$*,*

$$\mathcal{M}, c(x) \vDash \varphi \Leftrightarrow \mathcal{M}_{\approx}, x \vDash_{\approx} \varphi \Leftrightarrow \mathcal{M}^{\mathfrak{C}}, E^{-1}(x) \vDash^{\mathfrak{C}} \varphi$$

Now, we have all the background ingredients to introduce the BACs and show their application to Lewis counterfactuals.

2.2 Boolean Algebras of Conditionals in a Nutshell

In this section, we review the framework of Boolean algebras of conditionals (BACs) introduced by Flaminio, Godo, and Hosni (2020). In the first part, we focus on the technical properties of the BACs, such as their atomic structure, and the resulting conditional logics. In the second part, we draw some considerations on the philosophical impact of BACs on the debate over conditionals and their probability.

2.2.1 Definitions, Atomic Structures, and Other Key Properties

We assume that the reader is familiar with the basic properties of Boolean algebra (see for instance Davey and Priestley 2002 and Halmos and Givant 2009 for an extensive introduction to Boolean algebras). The framework of

BACs offers an innovative and privileged perspective on conditionals events: BACs are a valuable tool to analyze the algebraic properties of conditionals events, their logic, and their relation with probability measures. In what follows, we will review the definition and some basic properties of BACs from the work of Flaminio, Godo, and Hosni (2020).

Notation 2.4.

- *For an algebra* $\mathbf{X}$ *written in boldface,* X*, in italic, denotes the underlying set of* $\mathbf{X}$.

Definition 2.8 (Boolean Algebra of Conditionals, Flaminio, Godo, and Hosni 2020)**.** *Given any finite Boolean algebra* $\mathbf{A} = \langle A, \wedge, \vee, \neg, \top, \bot \rangle$*, consider:*

- $A' = A \setminus \{\bot\}$ *the set of all the elements of* A*, except for* $\bot$*;*
- $A \mid A' = \{(a \mid b) \mid a \in A, b \in A'\}$ *the set of all meaningful conditional events of the form* $(a \mid b)$*, read as "a given b" where b cannot be* $\bot$*;*
- *the free Boolean algebra* $\mathbf{Free}(A \mid A') = \langle Free(A \mid A'), \sqcap, \sqcup, \sim, \top^*, \bot^* \rangle$ *generated by* $A \mid A'$.

Then, consider the smallest congruence relation $\equiv_{\mathfrak{C}}$ *on* $\mathbf{Free}(A \mid A')$ *such that:*

(C1) $(b \mid b) \equiv_{\mathfrak{C}} \top^*$*, for all* $b \in A'$*;*

(C2) $(a_1 \mid b) \sqcap (a_2 \mid b) \equiv_{\mathfrak{C}} (a_1 \wedge a_2 \mid b)$*, for all* $a_1, a_2 \in A$*,* $b \in A'$*;*

(C3) $\sim(a \mid b) \equiv_{\mathfrak{C}} (\neg a \mid b)$*, for all* $a \in A$*,* $b \in A'$*;*

(C4) $(a \wedge b \mid b) \equiv_{\mathfrak{C}} (a \mid b)$*, for all* $a \in A$*,* $b \in A'$*;*

(C5) $(a \mid b) \sqcap (b \mid c) \equiv_{\mathfrak{C}} (a \mid c)$*, for all* $a \in A$*,* $b, c \in A'$ *such that* $a \leq b \leq c$.

Finally, the Boolean algebras of conditionals of $\mathbf{A}$*, denoted* $\mathcal{C}(\mathbf{A})$*, is defined as follows:*

$$\mathcal{C}(\mathbf{A}) = \mathbf{Free}(A \mid A')/_{\equiv_{\mathfrak{C}}}$$

To distinguish the operations of $\mathbf{A}$ *from those of* $\mathcal{C}(\mathbf{A})$*, the following signature is adopted:*

$$\mathcal{C}(\mathbf{A}) = (\mathcal{C}(A), \sqcap, \sqcup, \sim, \top_{\mathfrak{C}}, \bot_{\mathfrak{C}})$$

Notation 2.5.

Given a BAC, $\mathcal{C}(\mathbf{A}) = \langle \mathcal{C}(A), \sqcap, \sqcup, \sim, \top_{\mathfrak{C}}, \bot_{\mathfrak{C}} \rangle$, we adopt the following notation:

- *$\sqsubseteq$ is the natural order over $\mathcal{C}(\mathbf{A})$;*
- *Since $\mathcal{C}(\mathbf{A})$ is a quotient of $\mathbf{Free}(A \mid A')$, its elements are equivalence classes. However, without danger of confusion, we will henceforth identify an element $[t]_{\equiv_{\mathfrak{C}}}$ of $\mathcal{C}(\mathbf{A})$ with one of its representative elements, in particular, by t itself.*

Although Definition 2.8 may look heavy and intricate, the intuitive idea behind the BAC construction is rather simple. Fist of all, a BAC, $\mathcal{C}(\mathbf{A})$, is always constructed from a given finite Boolean algebra $\mathbf{A} = (A, \wedge, \vee, \neg, \bot, \top)$. Intuitively, if we interpret $\mathbf{A}$ as a space of (non-conditional) events like $a, b, c, \ldots$, its BAC, $\mathcal{C}(\mathbf{A})$, can be regarded as the space of *conditional events* of the form $(a \mid b), (a \mid c) \ldots$ read as "a given b", "a given c", ... Thus, a BAC $\mathcal{C}(\mathbf{A})$ includes objects of the form $(x \mid y)$ and their Boolean combinations. For instance, in $\mathcal{C}(\mathbf{A})$ we encounter objects of the form $(a \mid b) \sqcap (c \mid b)$, that is the conjunction of the objects $(a \mid b)$ and $(c \mid b)$, or $\neg(a \mid b)$, that is the complement of the object $(a \mid b)$, etc. However, these Boolean combinations of conditional objects, do not behave in a "wild" way inside $\mathcal{C}(\mathbf{A})$. Specifically, they obey the rules of probability. The congruence relation that induces the BAC is aimed to impose this probabilistic behavior to the wild conditional objects in $\mathbf{Free}(A \mid A')$. For instance, for every $b \in A'$, the conditional $(b \mid b)$ will be the top element of $\mathcal{C}(\mathbf{A})$, while $(\neg b \mid b)$ will be the bottom; this corresponds intuitively to the fact that the probability of $(b \mid b)$ is 1 and the probability of $(\neg b \mid b)$ is 0. In a conditional $(a \mid b)$ we can replace the consequent a by $a \wedge b$, that is, the conditionals $(a \mid b)$ and $(a \wedge b \mid b)$ are equal. This identity corresponds to the principle that the probability of $(a \wedge b \mid b)$ is equal to the probability of $(a \mid b)$. For all $a \in A$ and all $b \in A'$, $\sim (a \mid b) = (\neg a \mid b)$; this property corresponds to the fact that the probability of $(\neg a \mid b)$ is 1 minus the probability of $(a \mid b)$, i.e. $1 - P(a \mid b) = P(\neg a \mid b)$. More formally, some key properties of the objects in a BAC can be summarized as follows:

Lemma 2.6 (Flaminio, Godo, and Hosni 2020)**.** *Every BAC, $\mathcal{C}(\mathbf{A})$ is a Boolean algebra. Furthermore, the following identities hold in any BAC, $\mathcal{C}(\mathbf{A}) = \mathbf{Free}(A \mid A')/_{\equiv_{\mathfrak{C}}}$: for all $a, a' \in \mathbf{A}$ and $b, c \in \mathbf{A}'$:*

1. *$(b \mid b) = 1_{\mathfrak{C}}$;*
2. *$(a \mid b) \sqcap (c \mid b) = (a \wedge c \mid b)$;*
3. *$\sim(a \mid b) = (\neg a \mid b)$;*

4. $(a \wedge b \mid b) = (a \mid b)$;

5. if $a \leq b \leq c$, then $(a \mid b) \sqcap (b \mid c) = (a \mid c)$.

Some striking properties of a BAC emerge from the structure and characterization of its atoms. First, we need some notation:

Notation 2.6.

- *Given a finite set X, a permutation of the elements of X is a maximal sequence $\langle x_1, \ldots, x_n \rangle$ of pairwise different elements of X such that $\{x_1, \ldots, x_n\} = X$ and $x_1, \ldots, x_n \in X$. $Perm(X)$ is the set of all the permutations of the elements of X, i.e. $Perm(X) = \{\langle x_1, \ldots, x_n \rangle \mid \langle x_1, \ldots, x_n \rangle$ is a permutation of the elements of $X\}$*

- *consider a permutation $x = \langle x_1, \ldots, x_n \rangle \in Perm(X)$. for $1 \leq i \leq n$, $x[i]$ is the i-th projection of the permutation x, i.e. the i-th element appearing in x.*

Lemma 2.7 (Flaminio, Godo, and Hosni 2020, Theorem 4.4)**.** *Given a BAC, $C(\mathbf{A})$, the set of its atoms, $at(C(\mathbf{A}))$, is in one-to-one correspondence with the set of permutations of the atoms of $\mathbf{A}$. More formally:*

- $at(\mathbf{A}) = \{\alpha \mid \alpha$ *is an atom of* $\mathbf{A}\}$ *is the set of atoms of* $\mathbf{A}$;

- $at(C(\mathbf{A})) = \{\omega \mid \omega$ *is an atom of* $C(\mathbf{A})\}$ *is the set of atoms of* $C(\mathbf{A})$;

- $Perm(at(\mathbf{A})) = \{\langle \alpha_1, \ldots, \alpha_n \rangle \mid \langle \alpha_1, \ldots, \alpha_n \rangle$ *is a permutation of the elements of* $at(\mathbf{A})\}$ *is the set of permutations over* $at(\mathbf{A})$

There is a bijection $\Omega : Perm(at(\mathbf{A})) \rightarrowtail\!\!\!\!\!\twoheadrightarrow at(C(\mathbf{A}))$ between the set of permutations over $at(\mathbf{A})$ and the atoms of $C(\mathbf{A})$ such that:

$$\Omega(\langle \alpha_1, \ldots, \alpha_n \rangle) = (\alpha_1 \mid \top) \sqcap (\alpha_2 \mid \neg\alpha_1) \sqcap (\alpha_3 \mid \neg\alpha_1 \wedge \neg\alpha_2) \sqcap \cdots \sqcap (\alpha_n \mid \neg\alpha_1 \wedge \cdots \wedge \alpha_{n_1})$$

The last element involved in the $\sqcap$-meet operation above, i.e. $(\alpha_n \mid \neg\alpha_1 \wedge \cdots \wedge \alpha_{n_1})$, is equal to $\top_{\mathfrak{C}}$, hence it can be dropped.

As a corollary, we obtain some information on the cardinality of a BAC.

Corollary 2.2. *Given a BAC, $C(\mathbf{A})$, assume $|at(\mathbf{A})| = n$ for some $n \in \mathbb{N}$. Then, $|at(C(\mathbf{A}))| = n!$ and $C(\mathbf{A})$ has $2^{n!}$ elements.*

Another useful property is the following:

Lemma 2.8 (Flaminio, Godo, and Hosni 2020)**.** *Given a BAC, $\mathcal{C}(\mathbf{A})$, consider the subalgebra $(\mathbf{A} \mid \top)$ of $\mathcal{C}(\mathbf{A})$ whose underlying set is $\{(x \mid \top) \mid x \in A\}$, then $\mathbf{A}$ is isomorphic to $(\mathbf{A} \mid \top)$: $\mathbf{A} \cong (\mathbf{A} \mid \top)$*

Intuitively, the above lemma establishes that the original algebra $\mathbf{A}$ of non-conditional events is "contained" inside the algebra of conditionals events. Specifically, the (sub)algebra made of degenerated conditional events of the form $(x \mid \top)$ amounts to the same (up to isomorphism) original algebra.

The following example will help clarifying the BAC construction and its properties.

Example of a BAC

Assume our initial Boolean algebra $\mathbf{A}$ has 3 atoms, $at(\mathbf{A}) = \{\alpha_1, \alpha_2, \alpha_3\}$. The corresponding BAC, $\mathcal{C}(\mathbf{A})$, has 6 atoms, $at(\mathcal{C}(\mathbf{A})) = \{\omega_1, \omega_2, \dots, \omega_6\}$, and each of them, by Lemma 2.6, can be identified with a permutations of the elements α_1, α_2 and α_3. More explicitly:

$$
\begin{array}{lll}
\omega_1 \rightsquigarrow & \langle \alpha_1, \alpha_2, \alpha_3 \rangle \rightsquigarrow & (\alpha_1 \mid \top) \sqcap (\alpha_2 \mid \neg\alpha_1) \sqcap (\alpha_3 \mid \neg\alpha_1 \wedge \neg\alpha_2) \\
\omega_2 \rightsquigarrow & \langle \alpha_1, \alpha_3, \alpha_2 \rangle \rightsquigarrow & (\alpha_1 \mid \top) \sqcap (\alpha_3 \mid \neg\alpha_1) \sqcap (\alpha_2 \mid \neg\alpha_1 \wedge \neg\alpha_3) \\
\omega_3 \rightsquigarrow & \langle \alpha_2, \alpha_1, \alpha_3 \rangle \rightsquigarrow & (\alpha_2 \mid \top) \sqcap (\alpha_1 \mid \neg\alpha_2) \sqcap (\alpha_3 \mid \neg\alpha_2 \wedge \neg\alpha_2) \\
\omega_4 \rightsquigarrow & \langle \alpha_2, \alpha_3, \alpha_1 \rangle \rightsquigarrow & (\alpha_2 \mid \top) \sqcap (\alpha_3 \mid \neg\alpha_2) \sqcap (\alpha_1 \mid \neg\alpha_2 \wedge \neg\alpha_3) \\
\omega_5 \rightsquigarrow & \langle \alpha_3, \alpha_1, \alpha_2 \rangle \rightsquigarrow & (\alpha_3 \mid \top) \sqcap (\alpha_1 \mid \neg\alpha_3) \sqcap (\alpha_2 \mid \neg\alpha_3 \wedge \neg\alpha_1) \\
\omega_6 \rightsquigarrow & \langle \alpha_3, \alpha_2, \alpha_1 \rangle \rightsquigarrow & (\alpha_3 \mid \top) \sqcap (\alpha_2 \mid \neg\alpha_3) \sqcap (\alpha_1 \mid \neg\alpha_3 \wedge \neg\alpha_2)
\end{array}
$$

Each of the elements in light-gray is identical to $\top_{\mathfrak{C}}$, i.e. the top element of $\mathbf{A}$. Thus it can be dropped. This example is depicted in Figure 2.2

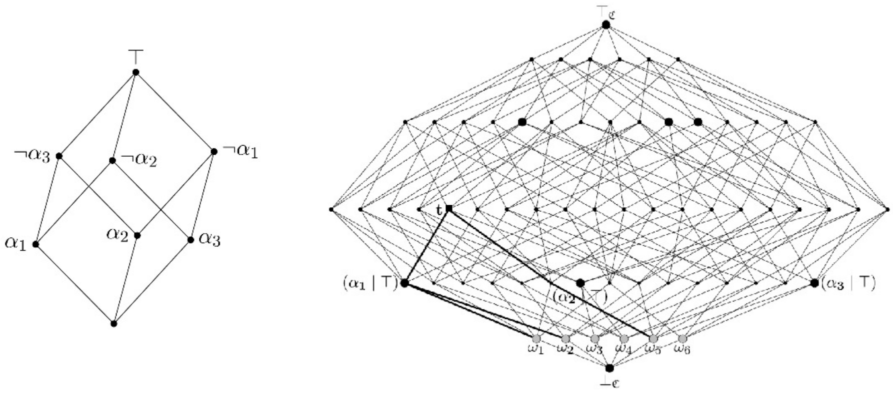


Figure 2.2: A graphical example of a BAC construction. Given a three-elements Boolean algebra (on the left), its corresponding BAC (on the right) contains 6 atoms and 64 elements. Moreover, the subalgebra $(\mathbf{A} \mid \top)$ of $C(\mathbf{A})$ is isomorphic to $\mathbf{A}$. The pictures are taken from (Flaminio, Godo, and Hosni 2020).

Notation 2.7.

By Lemma 2.6, with a slight abuse, we use the following notation:

- *we identify each atom ω of $C(\mathbf{A})$ with a permutation over $at(\mathbf{A})$, i.e. $\omega = \langle \alpha_1, \ldots, \alpha_n \rangle$ denotes an atom of $C(\mathbf{A})$;*
- *for $1 \leq i \leq |at(\mathbf{A})|$, and $\omega \in at(C(\mathbf{A}))$, $\omega[i]$ is the i-th projection of ω, i.e. $\omega[i]$ is the i-th element appearing in (the string identified with) ω.*

We have all the ingredients to characterize the relation between the atoms of a BAC and its elements:

Lemma 2.9 (Flaminio, Godo, and Hosni 2020, Proposition 4.7)**.** *Given a BAC, $C(\mathbf{A})$, let $\leq$ be the natural order over $\mathbf{A}$. For every atom $\omega \in at(C(\mathbf{A}))$, for every element of the form $(x \mid y)$ in $C(\mathbf{A})$,*

$$\omega \sqsubseteq (x \mid y) \Leftrightarrow \text{ there is a } i \text{ such that } \omega[i] \leq x \wedge y \text{ and for all } j < i, \omega[j] \nleq y$$

Namely, an atom ω of $C(\mathbf{A})$ is below a certain conditional $(x \mid y)$ iff the first element i appearing in ω that is below y in $\mathbf{A}$ ($i \leq y$) is such that it is below x too ($i \leq x$).

The following lemma recaps some results describing the relation among conditional objects in a BAC:

Lemma 2.10 (Flaminio, Godo, and Hosni 2020, Proposition 3.8)**.** *In every BAC, $C(\mathbf{A})$, the following hold for any $x, y, z \in A$:*

1. $(x \wedge y \mid \top) \sqsubseteq (x \mid y) \sqsubseteq (y \to x \mid \top)$;
2. *if* $x \leq y$ *then* $(x \mid z) \sqsubseteq (y \mid z)$;
3. $(z \mid x) \wedge (z \mid y) \sqsubseteq (z \mid x \vee y)$

Finally, we can introduce the semantics for the language $\mathcal{L}_{LBC}$. It will be useful to briefly recap some properties of the Lindenbaum construction.

Notation 2.8.

- $\dashv\vdash_{CPL} \subseteq For_{\mathcal{L}} \times For_{\mathcal{L}}$ *is the congruence relation of inter-derivability in classical logic between formulas in* $\mathcal{L}$;
- $\mathbf{L}$ *is the Lindenbaum algebra of CPL over the language* $\mathcal{L}$;
- $[]_{\dashv\vdash_{CPL}} : For_{\mathcal{L}} \to \mathbf{L}$ *is the canonical homomorphism of* $\mathcal{L}$ *into* $\mathbf{L}$ *mapping each formula in* $For_{\mathcal{L}}$ *to its* $\dashv\vdash_{CPL}$*-equivalence class in* $\mathbf{L}$*, i.e.* $[\varphi]_{\dashv\vdash_{CPL}} = \{\psi \in For_{\mathcal{L}} \mid \psi \dashv\vdash_{CPL} \varphi\}$.

Lemma 2.11. *Consider* $\mathbf{L}$*, i.e. the Lindenbaum algebra of CPL over the language* $\mathcal{L}$*. Recall that we are assuming that Var is finite. Then the following holds:*

1. *there is a one-to-one correspondence between atoms of* $\mathbf{L}$ *and classical valuations of* $\mathcal{L}$, $^* : Val_{CPL} \rightarrowtail\!\!\!\!\twoheadrightarrow at(\mathbf{L})$ *such that*

$$v* = [\bigwedge_{p:v(p)=1} p \wedge \bigwedge_{\neg p:v(p)=0} \neg p]_{\dashv\vdash_{CPL}}$$

2. *the bijection* $^* : Val_{CPL} \rightarrowtail\!\!\!\!\twoheadrightarrow at(\mathbf{L})$ *is such that for all classical valuations* v*, for all formulas* $\varphi \in For_{\mathcal{L}}$,

$$v(\varphi) = 1 \Leftrightarrow v* \leq [\varphi]_{\dashv\vdash_{CPL}}$$

where $\leq$ *is the natural order over* $\mathbf{L}$

Thus, for simplicity, we can identify classical valuation of $\mathcal{L}$ with atoms of $\mathbf{L}$.

The following lemma shows the relation between classical valuations/atoms of $\mathbf{L}$ with BACs:

Corollary 2.3 (Flaminio, Godo, and Hosni 2020). *$C(\mathbf{L})$ denotes the BAC obtained from $\mathbf{L}$. There is a bijection $\Omega^* : Perm(Val_{CPL}) \rightarrowtail\!\!\!\!\twoheadrightarrow at(C(\mathbf{L}))$ between atoms of $C(\mathbf{L})$ and permutations of classical valuations such that*

$$\Omega^*(\langle v_1, \ldots, v_n \rangle) = \Omega(\langle v_1^*, \ldots, v_n^* \rangle)$$

where * is the bijection from Lemma 2.11 and Ω is the bijection from Lemma 2.7.

Thus, for simplicity, we will thereby identify atoms of $C(\mathbf{L})$ with permutations of classical valuations of $\mathcal{L}$.

We have now all the ingredients to introduce the semantics for $\mathcal{L}_{LBC}$:

Definition 2.9. *Let $[]_{LBC} : For_{\mathcal{L}_{LBC}} \to C(\mathbf{L})$ be a homomorphism from formulas of $\mathcal{L}_{LBC}$ to $C(\mathbf{L})$ inductively defined as follows:*

- $[\neg\Phi]_{LBC} = \sim([\Phi]_{LBC})$
- $[\Phi \wedge \Psi]_{LBC} = [\Phi]_{LBC} \sqcap [\Psi]_{LBC}$
- $[\Phi \vee \Psi]_{LBC} = [\Phi]_{LBC} \sqcup [\Psi]_{LBC}$
- $[(\varphi \mid \psi)]_{LBC} = ([\varphi]_{\dashv\vdash_{CPL}} \mid [\psi]_{\dashv\vdash_{CPL}})$

A LBC-valuation is any element of $Perm(Val_{CPL})$, i.e. any permutation of classical valuations of $\mathcal{L}$. Furthermore, let $\vDash \subseteq Perm(Val_{CPL} \times For_{\mathcal{L}})$ be a relation between LBC-valuations $e = \langle v_1, \ldots, v_n \rangle$ and any formulas of $\mathcal{L}_{LBC}$ defined as follows: for all $e \in Perm(Val_{CPL})$, for all $\Phi \in For_{\mathcal{L}_{LBC}}$,

$$e \vDash \Phi \Leftrightarrow \Omega^*(e) \sqsubseteq [\Phi]_{LBC}$$

The following semantic clauses are readily provable:

Lemma 2.12. *For any LBC-valuation e, for all formulas $\Phi, \Psi \in For_{\mathcal{L}_{LBC}}$, the following holds:*

$$\begin{array}{lcl} e \vDash (\varphi \mid \psi) & \Leftrightarrow & \text{there is a } i \text{ such that } e[i](\varphi \wedge \psi) = 1 \text{ and for all } j < i, e[j](\psi) = 0 \\ e \vDash \Phi \wedge \Psi & \Leftrightarrow & e \vDash \Phi \text{ and } e \vDash \Psi \\ e \vDash \Phi \vee \Psi & \Leftrightarrow & e \vDash \Phi \text{ or } e \vDash \Psi \\ e \vDash \neg\Phi & \Leftrightarrow & e \nvDash \Phi \end{array}$$

Example of LBC-valuation

Assume $Var = p, q$, hence $Val_{CPL} = \{v_1, v_2, v_3, v_4\}$. Suppose:

- $v_1(p) = 1$ and $v_1(q) = 1$
- $v_2(p) = 1$ and $v_2(q) = 0$
- $v_3(p) = 0$ and $v_1(q) = 1$
- $v_4(p) = 0$ and $v_1(q) = 0$

Consider the *LBC*-valuation $e = \langle v_1, v_2, v_3, v_4 \rangle$. We have that:

- $e \vDash (p \mid q)$ since the first valuation appearing in e that makes q true ($v_1(q) = 1$) also makes p true ($v_1(p) = 1$);
- $e \nvDash (\neg p \mid \neg q)$ since the first valuation appearing in e that makes q false ($v_2(q) = 0$) does not make p false ($v_2(p) = 1$)

Finally, we adopt the following notation for logical consequence among formulas in $\mathcal{L}_{LBC}$:

Notation 2.9.

For $\Xi \cup \{\Phi\} \subseteq For_{\mathcal{L}_{LBC}}$,

$$\Xi \models_{LBC} \Phi \quad \Leftrightarrow \quad \begin{array}{l} \textit{for all LBC-valuations } e, \\ \textit{if } e \vDash \Psi \textit{ for all } \Psi \in \Xi, \textit{ then } e \vDash \Phi \end{array}$$

The logic induced by *LBC* logical consequence is studied in (Flaminio, Godo, and Hosni 2020) and corresponds to the logic $\vdash_{\mathbf{LBC}}$ induced by the following axiom system:

Axioms

- For any tautology of *CPL*, the formula resulting from a uniform replacement of the variables by conditionals in $\mathcal{L}_{LBC}$;
- $\vdash_{\mathbf{LBC}} (\varphi \mid \varphi)$
- $\vdash_{\mathbf{LBC}} \neg(\varphi \mid \psi) \leftrightarrow (\neg\varphi \mid \psi)$
- $\vdash_{\mathbf{LBC}} ((\varphi \mid \psi) \wedge (\delta \mid \psi)) \leftrightarrow (\varphi \wedge \delta \mid \psi)$

- $\vdash_{\mathbf{LBC}} (\varphi \mid \psi) \leftrightarrow (\varphi \wedge \psi \mid \psi)$
- $\vdash_{\mathbf{LBC}} (\varphi \mid \psi) \leftrightarrow (\varphi \mid \delta) \wedge (\delta \mid \psi)$, if $\vdash_{CPL} \varphi \supset \delta$ and $\vdash_{CPL} \delta \supset \psi$

Rules

- from $\vdash_{CPL} \varphi \supset \psi$ derive $\vdash_{\mathbf{LBC}} (\varphi \mid \delta) \supset (\varphi \mid \delta)$
- from $\vdash_{CPL} \delta \leftrightarrow \psi$ derive $\vdash_{\mathbf{LBC}} (\varphi \mid \delta) \leftrightarrow (\varphi \mid \psi)$
- Modus Ponens: $\Phi, \Phi \supset \Psi \vdash_{\mathbf{LBC}} \Psi$

2.2.2 A Short Story of the Motivation Behind the BACs

In this section, we draw some philosophical considerations concerning the motivations behind the framework of the BACs. Flaminio, Godo, and Hosni (2020) state that their work

> [...] contributes to a long-standing question [...] whose general form can be roughly stated as follows: *conditional probability is the probability of conditionals*.

Namely, inside a BAC, the probability of (the proposition expressed by) a conditional $(a \mid b)$ is indeed equal, under suitable assumptions on the probability distribution, to the conditional probability of a given b. Indeed, as we anticipated in the introduction, the BACs can be seen as a (successful) attempt to show that the probability of a conditional does correspond to its conditional probability. Consequently, the BACs can serve as a response to the triviality result and suggest a "conciliation" between the suppositional and propositional view on conditionals mentioned in the introduction. Let us start by reviewing the striking triviality result due to Hájek (1989):

Theorem 2.3 (Hájek 1989)**.** *If K is a finite set with cardinality greater than two, F is a field of subsets of K, and P a probability function defined on F, then there is no binary operation $\rightarrow$ on F such that for all $A, B \in F$, if $P(A) > 0$, then $P(A \rightarrow B) = P(B \mid A)$*

Before reviewing the above theorem, some clarifications are needed. First of all, the definition of a probability function over a (finite) Boolean algebra:

Definition 2.10. *Given a finite Boolean algebra $\mathbf{A} = \langle A, \wedge, \vee, \neg, \top, \bot \rangle$, a probability P over $\mathbf{A}$ is a function $P : \mathbf{A} \rightarrow [0, 1]$ from $\mathbf{A}$ to the real interval $[0, 1]$ satisfying the following conditions:*

1. $P(\top) = 1$
2. *if* $a \wedge b = \bot$, *then* $P(a \vee b) = P(a) + P(b)$

We say that P is positive when for all $\alpha \in at(\mathbf{A})$, $P(\alpha) > 0$. Furthermore, given a distribution P over $at(\mathbf{A})$, $P : at(\mathbf{A}) \to [0,1]$ such that $\sum_{\alpha \in at(\mathbf{A})} P(\alpha) = 1$, P can be extended to a probability $P : \mathbf{A} \to [0,1]$ over $\mathbf{A}$ as follows:

$$P(a) = \sum_{\alpha \leq a} P(\alpha)$$

We say that a distribution P is positive if for all $\alpha \in at(\mathbf{A})$, $P(\alpha) > 0$

As a consequence of the above definition, we obtain the following:

Lemma 2.13. *Given a non empty set of possible worlds W, $\langle \wp(W), \cap, \cup, ^{-}, \emptyset, W \rangle$ is a Boolean algebra where:*

- *$\cap$ is the meet operation;*
- *$\cup$ is the join operation*
- *$^{-}$ (set-theoretic complement) is the negation operation*
- *W is the top element*
- *$\emptyset$ is the bottom element*
- *$at(\wp(W)) = W$*

Hence, a distribution P over the elements of W, $P : W \to [0,1]$ such that $\sum_{w \in W} P(w) = 1$ can be extended to a probability over $\wp(W)$ as follows: for $X \subseteq W$

$$P(X) = \sum_{w \in X} P(w)$$

Since a proposition is a subset of $X \subseteq W$, $P(X)$ is the probability of the proposition X.

In light of this definition, Theorem 2.3 has the following consequence:

Corollary 2.4. *Consider a finite Boolean algebra $\mathbf{A} = \langle A, \wedge, \vee, \neg, \top, \bot \rangle$ with more than 4 elements and a positive probability distribution P over $at(\mathbf{A})$, there is no binary operation $\to$: $A \times A \to A$ definable in $\mathbf{A}$ such that for any two elements $a, b \in A$ with $P(b) > 0$, $P(a \to b) = P(b \mid a) = \frac{P(a \wedge b)}{P(a)}$*

The idea behind the Theorem 2.3 relies on a cardinality argument. Roughly, given a Boolean algebra $\mathbf{A}$ with more than 4 elements, *there are more conditional probabilities expressible over $\mathbf{A}$ than elements in A*, hence there are two elements $x, y \in A$ such that $P(x \mid y)$ cannot coincide with the probability of any element in A. As a consequence, any binary operation $\to: A \times A \to A$ would not capture all the conditional probabilities on A, namely there would be some two elements $x, y \in A$ such that $P(x \to y) \neq P(y \mid x)$.

The natural question arises of whether it would be possible to *expand* the number of element in $\mathbf{A}$ so to account for all the conditional probabilities expressible over $\mathbf{A}$. The BAC of $\mathbf{A}$ provides a positive answer to this question. Specifically, the BAC $\mathcal{C}(\mathbf{A})$ contains a sufficient number of elements to account for the conditional probabilities expressible over $\mathbf{A}$. Specifically, for any $x, y \in \mathbf{A}$ with $P(y) > 0$, the element $(x \mid y)$ would be an element of the BAC, $\mathcal{C}(\mathbf{A})$.

Observe that $\mathcal{C}(\mathbf{A})$ is itself a Boolean algebra, hence it is possible to define a probability distribution P over $at(\mathcal{C}(\mathbf{A}))$ that extends to a probability function over all the elements in $\mathcal{C}(\mathbf{A})$. We also observed that $\mathcal{C}(\mathbf{A})$ contains elements of the form $(x \mid y)$ with $x, y \in A$. Therefore, a probability distribution over $at(\mathcal{C}(\mathbf{A}))$ also assigns a probability to $(x \mid y)$, i.e. $P(x \mid y)$. Does $P(x \mid y)$ coincide with the conditional probability of x given y? The answer is "yes", but under suitable assumptions, as the following results will show. First, recall that an atom $\omega \in at(\mathcal{C}(\mathbf{A}))$ can be identified with a permutation of the elements of $at(\mathbf{A})$, and has the following form:

$$\omega = \Omega(\langle \alpha_1, \ldots, \alpha_n \rangle) = (\alpha_1 \mid \top) \sqcap (\alpha_2 \mid \neg\alpha_1) \sqcap (\alpha_3 \mid \neg\alpha_1 \wedge \neg\alpha_2) \sqcap \cdots \sqcap (\alpha_n \mid \neg\alpha_1 \wedge \cdots \wedge \alpha_{n_1})$$

Definition 2.11. *Consider a finite Boolean algebra* $\mathbf{A}$ *and any positive probability distribution* P *over* $at(\mathbf{A})$*;* P *can be canonically extended to a probability distribution* $\mu_P : at(\mathcal{C}(\mathbf{A})) \to [0, 1]$ *in the following way: for any* $\omega \in at(\mathcal{C}(\mathbf{A}))$*, assume* $\omega = (\alpha_1 \mid \top) \sqcap (\alpha_2 \mid \neg\alpha_1) \sqcap (\alpha_3 \mid \neg\alpha_1 \wedge \neg\alpha_2) \sqcap \cdots \sqcap (\alpha_n \mid \neg\alpha_1 \wedge \cdots \wedge \alpha_{n_1})$ *then*

$$\mu_P(\omega) = P(\alpha_1) \times \frac{P(\alpha_2 \wedge \neg\alpha_1)}{P(\neg\alpha_1)} \times \frac{P(\alpha_3 \wedge \neg\alpha_1 \wedge \neg\alpha_2)}{P(\neg\alpha_1 \wedge \neg\alpha_2)} \times \cdots \times \frac{P(\alpha_n \wedge \neg\alpha_1 \wedge \cdots \wedge \alpha_{n_1})}{P(\neg\alpha_1 \wedge \cdots \wedge \alpha_{n_1})}$$

The following theorem holds:

Theorem 2.4 (Flaminio, Godo, and Hosni 2020)**.** *Consider a finite Boolean algebra* $\mathbf{A}$ *and any positive probability distribution* P *over* $at(\mathbf{A})$*. Consider the corresponding BAC* $\mathcal{C}(\mathbf{A})$ *and the canonically extended probability* μ_P *according to Definition 2.11. For all* $a, b \in A$ *with* $P(b) > 0$*, the following holds:*

$$\mu_P(a \mid b) = \sum_{\omega \in at(\mathcal{C}(\mathbf{A})) : \omega \sqsubseteq (a|b)} \mu_P(\omega) = P(a \mid b) = \frac{P(a \wedge b)}{P(b)}$$

and

$$\mu_P(a \mid \top) = P(a)$$

The above theorem establishes that the BAC, $\mathcal{C}(\mathbf{A})$, "include" all the probabilities (conditional and non conditionals) that can be expressed over $\mathbf{A}$. Furthermore, one could equip $\mathcal{C}(\mathbf{A})$ with a partial operation $\rightarrow$: $(A \mid \top) \times (A \mid \top) \rightarrow \mathcal{C}(\mathbf{A})$ such that for any two elements of the form $(a \mid \top), (b \mid \top)$ with $P(a) > 0$, $(a \mid \top) \rightarrow (b \mid \top) = (b \mid a)$. In this sense, in the BAC $\mathcal{C}(\mathbf{A})$, we could define a binary conditional whose probability coincides with the corresponding conditional probability, in fact:

$$\mu_P((a \mid \top) \rightarrow (b \mid \top)) = \mu_P(b \mid a) = \frac{P(b \wedge a)}{P(a)}$$

Let us try to apply the above considerations to the setting of possible worlds. Consider a set W of possible worlds and all the propositions definable over W, i.e. the powerset Boolean algebra $\wp(\mathbf{W}) = \langle \wp(W), \cup, \cap, ^{-}, W, \emptyset \rangle$. Then consider a positive probability distribution over W. Theorem 2.3 implies that there is no binary operation $\rightarrow$: $\wp(W) \times \wp(W) \rightarrow \wp(W)$ over the algebra of propositions of W such that for any two propositions $X, Y \in \wp(W)$ with $P(Y) > 0$, $P(X \rightarrow Y) = \sum_{w \in X \rightarrow Y} P(w) = \frac{P(X \cap Y)}{P(X)} = P(Y \mid X)$. However, under a suitable expansion of our algebra of propositions induced by the BAC construction, we can retrieve the equation between probability of conditionals and conditional probabilities. In particular, consider the BAC of $\wp(\mathbf{W})$, i.e. $\mathcal{C}(\wp(\mathbf{W}))$; the atoms of $\mathcal{C}(\wp(\mathbf{W}))$ would constitute our "new possible worlds". In particular, $\mathcal{C}(\wp(\mathbf{W}))$, being a Boolean algebra, would be isomorphic to $\langle \wp(at(\mathcal{C}(\wp(\mathbf{W})))), \cap, \cup, ^{-}, at(\mathcal{C}(\wp(\mathbf{W}))), \emptyset \rangle$. Therefore, for $X, Y \in \wp(W)$ with $P(Y) > 0$, the proposition expressed by $(X \mid Y)$ is the set of the new possible worlds, i.e. the atoms of $\mathcal{C}(\wp(\mathbf{W}))$, that are below the element $(X \mid Y)$ in $\mathcal{C}(\wp(\mathbf{W}))$. Now, if we consider the canonical extension of P over $at(\wp(\mathbf{W}))$, i.e. μ_P, we obtain that the probability of the proposition expressed by $(X \mid Y)$ coincides with its corresponding conditional probability:

$$\mu_P(Y \mid X) = P(Y \mid X) = \frac{P(X \cap Y)}{P(Y)}$$

and moreover

$$\mu_P(X \mid W) = P(X)$$

To recap, given an algebra of proposition $\wp(W)$ and a positive probability distribution P over W, the BAC construction allows us to expand our initial algebra of propositions together with the probability distribution P in such a way that:

- any original proposition X can be identified with the new propositions $(X \mid W)$,

- the proposition $(X \mid Y)$ is assigned the conditional probability of X given Y. In particular, the probability of the proposition expressed by $(X \mid Y)$, coincide with the corresponding conditional probability, under μ_P:

$$\mu_P(X \mid Y) = P(X \mid Y) = \frac{P(X \cap Y)}{P(Y)}$$

Therefore, in the BAC, we can identify a proposition $(X \mid Y)$ whose probability under μ_P coincides with its corresponding conditional probability under P. In this sense, the BAC framework mitigates the consequences of the triviality result by showing that it is always possible (in the finite setting) to construct an algebra of propositions in which conditional probabilities can be expressed as probabilities of specific propositions.

2.3 Modal Boolean Algebras of Conditionals

In this section, we introduce a special modal extension of the framework of BACs. Specifically, we are going to define new algebraic structures obtained by adding a modal operator $\Box$ to a BAC, so as to have *modal* conditional objects of the form $\Box(x \mid y)$. We will demonstrate how these novel modal structures serve as algebraic models for Lewis counterfactual conditionals, by showing that a counterfactual conditional of the form $\varphi \mathrel{\Box\!\!\rightarrow} \psi$ can be interpreted as the modal conditional $\Box(\psi \mid \varphi)$.

2.3.1 Definitions and Basic Properties

The fundamental structure of our novel framework will be a special kind of Modal BAC that we call *lewis algebra*:

Definition 2.12. *A Modal BAC, $\langle C(\mathbf{A}), \Box\rangle$ consists of a BAC equipped with a normal unary operator satisfying the following constraints: for all x, y in $C(\mathbf{A})$*

(L0) $\Box\top_{\mathfrak{C}} = \top_{\mathfrak{C}}$

(L1) $\Box(x \wedge y) = \Box x \wedge \Box y$

Furthermore, we can define the modal operator $\Diamond$, dual to $\Box$, as customary: for all x in $C(\mathbf{A})$, $\Diamond x = \neg\Box\neg x$.

A lewis algebra is a modal BAC, $\langle C(\mathbf{A}), \Box\rangle$, where $\Box$ satisfies the following additional constraints: for all a, b, c in $\mathbf{A}$

(L2) $\Box(a \mid \top) = (a \mid \top)$

(L3) $\Box(a \mid a \vee b) \sqcup \Box(b \mid a \vee b) \sqcup (\Box(c \mid a \vee b) \Leftrightarrow (\Box(c \mid a) \sqcap \Box(c \mid b))) = \top_{\mathfrak{C}}$

Clearly, to any finite algebra $\mathbf{A}$ there correspond possibly several, not isomorphic, Lewis algebras depending on how many $\Box$'s operators satisfying $(L1)$, $(L2)$ and $(L3)$ can be defined upon the conditional algebra $\mathcal{C}(\mathbf{A})$. Since every $\mathcal{C}(\mathbf{A})$ is a Boolean algebra, every lewis algebra $\langle \mathcal{C}(\mathbf{A}), \Box \rangle$ is a *Boolean algebra with operators* (see Blackburn, de Rijke, and Venema 2001; Jipsen 1992).

Let us show some basic properties of the elements in a lewis algebra:

Lemma 2.14. *The following equations hold in every Lewis algebra* $\langle \mathcal{C}(\mathbf{A}), \Box \rangle$*:*

1. $\Box(a \mid a) = \top_{\mathfrak{C}}$;
2. $(a \wedge b \mid \top) \sqsubseteq \Box(a \mid b) \sqsubseteq (b \rightarrow a \mid \top)$;
3. $\Diamond(a \mid \top) = (a \mid \top)$;
4. $\Box(a \mid b) \sqsubseteq \Diamond(a \mid b)$;
5. $\Box((c \mid a) \wedge (c \mid b)) \sqsubseteq \Box(c \mid a \vee b)$.

Proof. Let us start by noticing that equation $(L1)$ implies, as customary, that $\Box$ is monotone. That is to say, for all $t, s \in \langle \mathcal{C}(\mathbf{A}), \Box \rangle$, if $t \sqsubseteq s$, then $\Box t \sqsubseteq \Box s$.

1. By Lemma 2.6-1. and Definition 2.12-$(L0)$.
2. By Lemma 2.10-1., $(a \wedge b \mid \top) \sqsubseteq (a \mid b) \sqsubseteq (b \rightarrow a \mid \top)$. Then, by monotonicity of $\Box$,$\Box(a \wedge b \mid \top) \sqsubseteq \Box(a \mid b) \sqsubseteq \Box(b \rightarrow a \mid \top)$. Thus, Definition 2.12-$(L2)$ implies $(a \wedge b \mid \top) \sqsubseteq \Box(a \mid b) \leq (b \rightarrow a \mid \top)$.
3. By Definition 2.12-$(L2)$, $\Diamond(a \mid \top) =\sim \Box \sim (a \mid \top) =\sim \Box(\neg a \mid \top) =\sim (\neg a \mid \top) = (\neg\neg a \mid \top) = (a \mid \top)$.
4. $\Box(a \mid b) \Rightarrow \Diamond(a \mid b) =\sim \Box(a \mid b) \sqcup \Diamond(a \mid b) = \Diamond \sim (a \mid b) \sqcup \Diamond(a \mid b) = \Diamond(\neg a \mid b) \sqcup \Diamond(a \mid b) = \Diamond(\neg a \vee a \mid b) = \Diamond(\top \mid b) = \Diamond(\top \mid \top) = (\top \mid \top)$, where the last equality follows from point 3 above. Thus, $\Box(a \mid b) \sqsubseteq \Diamond(a \mid b)$.
5. By Lemma 2.10-3. and monotonicity of $\Box$.

□

Lewis algebras represent a special class of Boolean algebras with operators. Hence, we can apply the general theory of Jónsson-Tarski duality and analyze the Kripke frames associated to Lewis algebras. The following lemma is a straightforward consequence of the Jónsson-Tarski duality:

Lemma 2.15 (Jónsson-Tarski duality, Blackburn, de Rijke, and Venema 2001; Ono 2019)**.** *Each Modal BAC,* $\langle \mathcal{C}(\mathbf{A}), \Box \rangle$*, being a modal algebra, induces a Kripke frame* $\langle at(\mathcal{C}(\mathbf{A})), R \rangle$ *where:*

- *$R \subseteq at(C(\mathbf{A})) \times at(C(\mathbf{A}))$ is a binary accessibility relation over* $\mathrm{at}(C(\mathbf{A}))$ *defined as follows: for all $\omega, \omega' \in at(C(\mathbf{A}))$,*

$$\omega R \omega' \Leftrightarrow \textit{ for all } x \textit{ in } C(\mathbf{A}), \textit{ if } \omega \sqsubseteq \Box x \textit{ then } \omega' \sqsubseteq x$$

We call "Lewis frame" the dual Kripke frame of a Lewis algebra.

Furthermore, we introduce a relation $\models \subseteq at(C(\mathbf{A})) \times C(\mathbf{A})$ between atoms of $C(\mathbf{A})$ and elements of $C(\mathbf{A})$ defined as follows: for all $\omega \in at(C(\mathbf{A}))$, for all $x \in C(A)$,

$$\omega \models x \Leftrightarrow \omega \sqsubseteq x$$

Thus, the following is readily provable:

$$\begin{array}{lcl} \omega \models (a \mid b) & \Leftrightarrow & \textit{there is a } i \textit{ such that } \omega[i] \leq a \wedge b \textit{ and for all } j < i, \omega[j] \nleq b \\ \omega \models x \sqcap y & \Leftrightarrow & \omega \models x \textit{ and } \omega \models y \\ \omega \models x \sqcup y & \Leftrightarrow & \omega \models x \textit{ or } \omega \models y \\ \omega \models \sim x & \Leftrightarrow & \omega \not\models x \\ \omega \models \Box x & \Leftrightarrow & \textit{for all } \omega' \textit{ such that } \omega R \omega', \omega' \models x \end{array}$$

The Jónsson-Tarski duality implies that conditions on the modal operator $\Box$ in a modal algebra $\langle \mathbf{X}, \Box \rangle$ mirror conditions over the accessibility relation in the dual Kripke frame $\langle at(\mathbf{X}), R \rangle$. For instance, if $\Box$ satisfies the following requirement $\Box x \leq x$ in $\langle \mathbf{X}, \Box \rangle$, then the relation R in $\langle at(\mathbf{X}), R \rangle$ will be reflexive, i.e. for all $\alpha \in at(\mathbf{A})$, $\alpha R \alpha$. More formally

$$\Box x \leq x \text{ holds for all } x \text{ in } \langle \mathbf{X}, \Box \rangle \Leftrightarrow R \text{ is reflexive in } \langle at(\mathbf{X}), R \rangle$$

Thus, the question about what kind of constraints hold in a Lewis frame naturally arises. Specifically, we will analyze what kind of conditions the properties of $\Box$ in $\langle C(\mathbf{A}), \Box \rangle$ force on the relation R in $\langle at(C(\mathbf{A})), R \rangle$, in particular properties $(L2)$ and $(L3)$ in Definition 2.12. Some notation will be useful:

Notation 2.10.

- *for a Boolean algebra $\mathbf{A}$ and an element x in $\mathbf{A}$, let $\leq$ be the natural order over $\mathbf{S}$, $[x] = \{\alpha \in at(\mathbf{A}) \mid \alpha \leq x\}$ denotes the set of atoms below x;*
- *for a Kripke frame $\langle W, R \rangle$ and a $w \in W$, $R[w] = \{v \in W \mid wRv\}$ denotes the set of accessible worlds from w.*

Now, we can introduce the following tool:

Definition 2.13 (Selection Function). *Given a Lewis frame $\langle at(\mathcal{C}(\mathbf{A})), R\rangle$, a selection function is a map $f : A \times \mathrm{at}(\mathcal{C}(\mathbf{A})) \to \wp(\mathrm{at}(\mathbf{A}))$ defined as follows: for all $a \in A$, for all $\omega \in \mathrm{at}(\mathcal{C}(\mathbf{A}))$,*

$$f(a, \omega) = \{\omega'[i] \in at(\mathbf{A}) \mid \omega R \omega', \omega'[i] \leq a, \text{ and for all } j < i, \omega'[j] \nleq a\}$$

Intuitively $f(a, \omega)$ selects, for all accessible ω' from ω, the first element appearing in ω' that is below a. The selection function tool allows us characterize in a more familiar way the atoms below objects of the form $\Box(a \mid b)$. Specifically, the following remark will be useful.

Remark 2.1. *Recall the relation $\vDash \subseteq at(\mathcal{C}(\mathbf{A})) \times \mathcal{C}(\mathbf{A})$ in Lemma 2.15. By Definition 2.13 and Lemma 2.15, we have that for all $\omega \in at(\mathcal{C}(\mathbf{A}))$, $\omega \sqsubseteq \Box(a \mid b) \Leftrightarrow f(b, \omega) \subseteq [a]$. Thus:*

$$\omega \vDash \Box(a \mid b) \Leftrightarrow f(b, \omega) \subseteq [a]$$

Proof. Readily follows from Definition 2.13 and Lemma 2.15. □

At this point, it is worth noticing how the conditions for an atom to be below $\Box(a \mid b)$ resemble the semantics conditions for counterfactuals according to Definition 0.3. Hence, the reader may already expect that the object $\Box(a \mid b)$ interprets the counterfactual $a \Box\!\!\rightarrow b$. However, before demonstrating the main result, we need to examine the properties of Lewis frames. The properties on a Lewis frame $\langle \mathcal{C}(\mathbf{A}), R\rangle$ and the corresponding conditions on $\langle \mathcal{C}(\mathbf{A}), \Box\rangle$ are summarized in the following Table 2.3.1

	Dual Kripke Frame: $\langle at(\mathcal{C}(\mathbf{A})), R\rangle$	
	Conditions on R	Conditions on Modal BAC: $\langle \mathcal{C}(\mathbf{A}), R\rangle$
(Ser) Seriality	for all $\omega \in at(\mathcal{C}(\mathbf{A}))$, there is a ω' such that $\omega R \omega'$	$\Box(a \mid \top) = (a \mid \top)$
(Cen) Centering	for all $\omega, \omega' \in at(\mathcal{C}(\mathbf{A}))$, if $\omega R \omega'$ then $\omega[1] = \omega[1]$	
(Sph) Sphericality	for all $a, b \in A$, for all $\omega \in at(\mathcal{C}(\mathbf{A}))$, either $f(a \vee b, \omega) \subseteq [a]$ or $f(a \vee b, \omega) \subseteq [b]$ or $f(a \vee b, \omega) = f(a, \omega) \cup f(b, \omega)$	$(L3)$ in Definition 2.12

Table 2.2: The table summarizes the properties of a dual Kripke frame $\langle \mathcal{C}(\mathbf{A}), R\rangle$ and the corresponding conditions on $\langle \mathcal{C}(\mathbf{A}), \Box\rangle$

More precisely, the following lemma holds:

Lemma 2.16. *Consider a modal BAC $\mathcal{C}(\mathbf{A}), \Box\rangle$ and its dual Kripke frame $\langle at(\mathcal{C}(\mathbf{A})), R\rangle$. The following hold:*

1. $\left.\begin{array}{l}\langle at(\mathcal{C}(\mathbf{A})), R\rangle \textit{ satisfies } \textbf{Ser} \textit{ in Table 2.3.1} \\ \langle at(\mathcal{C}(\mathbf{A})), R\rangle \textit{ satisfies } \textbf{Cen} \textit{ in Table 2.3.1}\end{array}\right\} \Leftrightarrow$ *(L2) in Definition 2.12 holds in $\langle \mathcal{C}(\mathbf{A}), \Box\rangle$*

2. *$\langle at(\mathcal{C}(\mathbf{A})), R\rangle$ satisfies* **Sph** *in Table 2.3.1* $\Leftrightarrow$ *(L3) in Definition 2.12 holds in $\langle \mathcal{C}(\mathbf{A}), \Box\rangle$*

Proof.

1. ($\Rightarrow$) Consider $\omega \in \text{at}(\mathcal{C}(\mathbf{A}))$ such that $\omega \sqsubseteq \Box(a \mid \top)$. Then, for all ω' such that $\omega R \omega'$, $\omega' \sqsubseteq (a \mid \top)$. Since, by **Ser**, $R[\omega] \neq \emptyset$, take a $\omega' \in R[\omega]$. By assumption, it holds that $\omega' \sqsubseteq (a \mid \top)$, thus $\omega'[1] \leq a$. By **Cen**, $\omega'[1] = \omega[1]$, hence it holds that $\omega[1] \leq a$. So, $\omega \sqsubseteq (a \mid \top)$, therefore $\Box(a \mid \top) \sqsubseteq (a \mid \top)$.

   Now, consider $\omega \in \text{at}(\mathcal{C}(\mathbf{A}))$ such that $\omega \sqsubseteq (a \mid \top)$. Then, $\omega[1] \leq a$. Since, by **Ser**, $R[\omega] \neq \emptyset$, take $\omega' \in R[\omega]$; since by **Cen** $\omega'[1] = \omega[1]$, then $\omega'[1] \leq a$. Therefore $\omega' \sqsubseteq (a \mid \top)$. Since ω was taken arbitrarily in $R[\omega]$, then for all $\omega^* \in R[\omega]$, $\omega^* \sqsubseteq (a \mid \top)$. Hence, $\omega \sqsubseteq \Box(a \mid \top)$, and so $(a \mid \top) \sqsubseteq \Box(a \mid \top)$.

   Therefore, $\Box(a \mid \top) = (a \mid \top)$

   ($\Leftarrow$) By contraposition, assume that the Lewis frame $\langle at(\mathcal{C}(\mathbf{A})), R\rangle$ does not satisfies **Ser**. Thus, there is a $\omega \in at(\mathcal{C}(\mathbf{A}))$ such that for all $\omega' \in at(\mathcal{C}(\mathbf{A}))$ it is not the case that $\omega R \omega'$. Furthermore, let $a \in A$ be such that $\omega[1] \nleq a$. Therefore, vacuously $\omega \vDash \Box(a \mid \top)$, since no ω' is accessible from ω. On the other hand, $\omega \nvDash (a \mid \top)$ since $\omega[1] \leq \top$, but $\omega[1] \nleq a$ by assumption. Therefore, $\Box(a \mid \top) \neq (a \mid \top)$.

   By contraposition, assume that the Lewis frame $\langle at(\mathcal{C}(\mathbf{A})), R\rangle$ does not satisfies **Cen**. Thus, there are $\omega, \omega' \in at(\mathcal{C}(\mathbf{A}))$ such that $\omega R \omega'$, and $\omega[1] \neq \omega'[1]$. Let $a \in A$ be such that $\omega[1] \leq a$, but $\omega'[1] \nleq a$ (for instance, take $a = \omega[1]$). Then, $\omega \vDash (a \mid \top)$, but $\omega \nvDash \Box(a \mid \top)$. Therefore $\Box(a \mid \top) \neq (a \mid \top)$.

2. ($\Rightarrow$) Consider any $\omega \in at(\mathcal{C}(\mathbf{A}))$ and any $a, b \in A$. By **Sph**, either $f(a \vee b, \omega) \subseteq [a]$ or $f(a \vee b, \omega) \subseteq [b]$ or $f(a \vee b, \omega) = f(a, \omega) \cup f(b, \omega)$. If $f(a \vee b, \omega) \subseteq [a]$ holds, then by Remark 2.1 and Lemma 2.15, $\omega \sqsubseteq \Box(a \mid a \vee b)$. Similarly, if $f(a \vee b, \omega) \subseteq [b]$ holds, then $\omega \sqsubseteq \Box(b \mid a \vee b)$. If $f(a \vee b, \omega) = f(a, \omega) \cup f(b, \omega)$, consider any $c \in A$. We have that $\omega \sqsubseteq \Box(a \mid a \vee b) \sqcup \Box(b \mid a \vee b) \sqcup (\Box(c \mid a \vee b) \Leftrightarrow (\Box(c \mid a) \sqcap \Box(c \mid b)))$. Hence, $\omega \sqsubseteq \Box(a \mid a \vee b) \sqcup \Box(b \mid a \vee b) \sqcup \Box(a \mid a \vee b) \sqcup \Box(b \mid a \vee b) \sqcup (\Box(c \mid a \vee b) \Leftrightarrow (\Box(c \mid a) \sqcap \Box(c \mid b)))$.

   Therefore $(L3)$ holds.

   ($\Leftarrow$) First, notice that for all $a, b \in A$, for all $\omega \in at(\mathcal{C}(\mathbf{A}))$, $f(a \vee b, \omega) \subseteq f(a, \omega) \cup f(b, \omega)$ always holds in every Lewis frame.

   By contraposition, assume that the Lewis frame $\langle at(\mathcal{C}(\mathbf{A})), R\rangle$ does not satisfies **Sph**. Thus, there are $\omega \in \text{at}(\mathcal{C}(\mathbf{A}))$ and $a, b \in \mathbf{A}$ such that $f(a \vee b, \omega) \nsubseteq [a]$ and $f(a \vee b, \omega) \nsubseteq [b]$ and $f(a, \omega) \cup f(b, \omega) \nsubseteq f(a \vee b, \omega)$. Then, by Remark 2.1, $\omega \nvDash \Box(a \mid a \vee b)$, $\omega \nvDash \Box(b \mid a \vee b)$. Moreover, since $f(a, \omega) \cup f(b, \omega) \nsubseteq f(a \vee b, \omega)$, we have

that for some $\alpha \in f(a,\omega) \cup f(b,\omega)$, $\alpha \notin f(a \vee b,\omega)$. Without loss of generality, assume $\alpha \in f(a,\omega)$. Now, consider $c = \bigvee(f(a \vee b,\omega))$, clearly $c \in A$ and $f(a \vee b,\omega) = [c]$. So, by Remark 2.1, $\omega \models \Box(c \mid a \vee b)$; however, since $\alpha \notin f(a \vee b,\omega)$, by assumption, then $\alpha \not\leq c$. And so, $\omega \not\models \Box(c \mid a)$, hence $\omega \not\models \Box(c \mid b) \wedge \Box(c \mid a)$, thus $\omega \not\models \Box(c \mid a \vee b) \Rightarrow \Box((c \mid b) \wedge (c \mid a))$.
Therefore, $(L3)$ does not hold in $\langle at(C(\mathbf{A})), R\rangle$.

□

The properties of **Ser** and **Cen** are transparent: **Ser** establish that each element must have a successor; **Cen** establishes that two accessible elements share the same initial element. For example, it cannot be the case that $\omega = \langle\alpha_1,\alpha_2,\alpha_3\rangle$ accesses to $\omega' = \langle\alpha_2,,\alpha_3,\alpha_1\rangle$ since the first element in ω is α_2, while α_1 is the first element appearing in ω_1. The following example will clarify the definition of a Lewis frame.

Example of a Lewis Frame

Consider the BAC from Figure 2.2, and consider a modal operator □ defined over it such that □ satisfies conditions $(L1)--(L3)$ from Definition 2.12, so as to obtain a Lewis algebra. The Lewis frames associated to it may have the following form: $\langle at(C(\mathbf{A})), R\rangle$ where:

- $at(C(\mathbf{A})) =$
  $= \{\langle\alpha_1,\alpha_2,\alpha_3\rangle, \langle\alpha_1,\alpha_3,\alpha_2\rangle, \langle\alpha_2,\alpha_1,\alpha_3\rangle, \langle\alpha_2,\alpha_3,\alpha_1\rangle, \langle\alpha_3,\alpha_1,\alpha_2\rangle, \langle\alpha_3,\alpha_2,\alpha_1\rangle\}$
- R is such that:
  - $R[\langle\alpha_1,\alpha_2,\alpha_3\rangle] = \{\langle\alpha_1,\alpha_2,\alpha_3\rangle, \langle\alpha_1,\alpha_3,\alpha_1\rangle\}$
  - $R[\langle\alpha_1,\alpha_3,\alpha_2\rangle] = \{\langle\alpha_1,\alpha_3,\alpha_2\rangle\}$
  - $R[\langle\alpha_2,\alpha_1,\alpha_3\rangle] = \{\langle\alpha_2,\alpha_1,\alpha_3\rangle, \langle\alpha_2,\alpha_3,\alpha_1\rangle\}$
  - $R[\langle\alpha_2,\alpha_3,\alpha_1\rangle] = \{\langle\alpha_2,\alpha_3,\alpha_1\rangle\}$
  - $R[\langle\alpha_3,\alpha_1,\alpha_2\rangle] = \{\langle\alpha_3,\alpha_1,\alpha_2\rangle, \langle\alpha_3,\alpha_2,\alpha_1\rangle\}$
  - $R[\langle\alpha_3,\alpha_2,\alpha_1\rangle] = \{\langle\alpha_3,\alpha_2,\alpha_1\rangle\}$

R is such that each atom accesses to some other atom, and moreover each atom only accesses to atoms having its same initial element. In a picture:

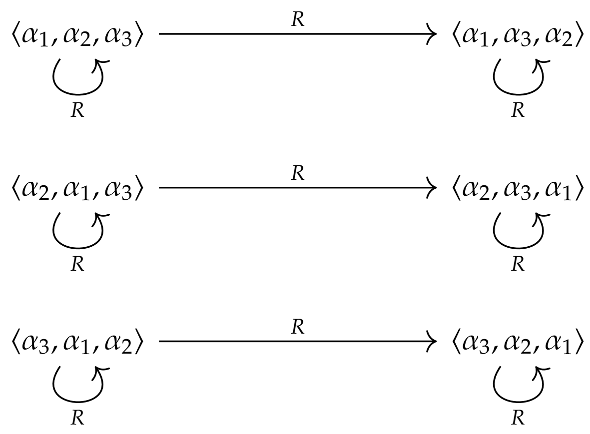


Figure 2.3: A picture of a Lewis frame

Furthermore, it is easy to see the the following additional properties of the selection function hold:

Lemma 2.17. *For every Lewis frame $\langle C(\mathbf{A}), \Box \rangle$, consider the selection function f defined over it. Then the following properties hold: for all $\omega \in at(C(\mathbf{A}))$ and all $a, b \in A$,*

1. $f(a, \omega) \subseteq [a]$;
2. *if* $f(a, \omega) \subseteq [b]$ *and* $f(b, \omega) \subseteq [a]$, *then* $f(a, \omega) = f(b, \omega)$;
3. *if* $\omega[1] \leq a$, *then* $f(a, \omega) = \{\omega[1]\}$

Proof.

1. Immediately follows from Definition 2.13 of selection function;
2. Assume *ad absurdum* that $f(a, \omega) \neq f(b, \omega)$ but $f(a, \omega) \subseteq [b]$ and $f(b, \omega) \subseteq [a]$. Thus, without loss of generality, there is a $\alpha \in f(a, \omega)$ such that $\alpha \notin f(b, \omega)$ by definition of f, for some $\omega' \in R[\omega]$, α is the first element appearing in ω' such that $\alpha \leq a$. By assumption, $\alpha \leq b$ too, since $f(b, \omega) \subseteq [b]$. Furthermore, $\alpha \notin f(b, \omega)$; this means that for all $\omega' \in R[\omega']$ there is a β such that $\beta \neq \alpha$, $\beta \leq b$, and β appears earlier than α in ω. By assumption, $f(b, \omega) \subseteq [a]$. Thus, since all such β's appear earlier than α, it cannot be the case that $\alpha \in f(a, \omega)$, contradicting our assumption.
3. Immediately follows from **Cen** and **Ser**.

□

The properties of the selection function introduced above in Lemma 2.17, together with $(L3)$ from Definition 2.12, strictly resemble the properties of the selection function in those functional Lewisian models satisfying **C**, i.e. **VC** models. Hence, we are getting closer to see how Lewis counterfactuals, i.e. conditionals obeying the axioms and rules of the logic **VC**, can be interpreted into conditional objects of the form $\Box(a \mid b)$ inside a Lewis algebra. But first, a more transparent characterization of the property **Sph** is required. Indeed, we have showed how **Sph** predicates a certain behavior of the selection function defined over a Lewis frame. However, the selection function is a derived tool, which is defined starting from the accessibility relation and the structure of the elements of a Lewis frame. Hence, at a deeper level, the property **Sph** must characterize certain features of the accessibility relation and the structure of the atoms of a Lewis frame. In fact, in the next section, we will better clarify the impact of **Sph** on the structure of a Lewis frame.

2.3.2 Sphericality

In order to provide a more transparent characterization of the condition imposed by **Sph**, we need a complicated construction. First, recall once more that a Lewis frame $\langle at(C(\mathbf{A})), R\rangle$ is a Kripke frames whose elements are atoms of a BAC; hence, these elements, by Lemma 2.7, can be identified with permutations of the atoms of **A**. Furthermore, $R[\omega] \neq \emptyset$ by **Ser** and every $\omega' \in R[\omega]$ is such that $\omega[1] = \omega'[1]$, by **Cen** i.e., ω and all its accessible atoms share the its first element. Then, $R[\omega]$ can be arranged as the following figure:

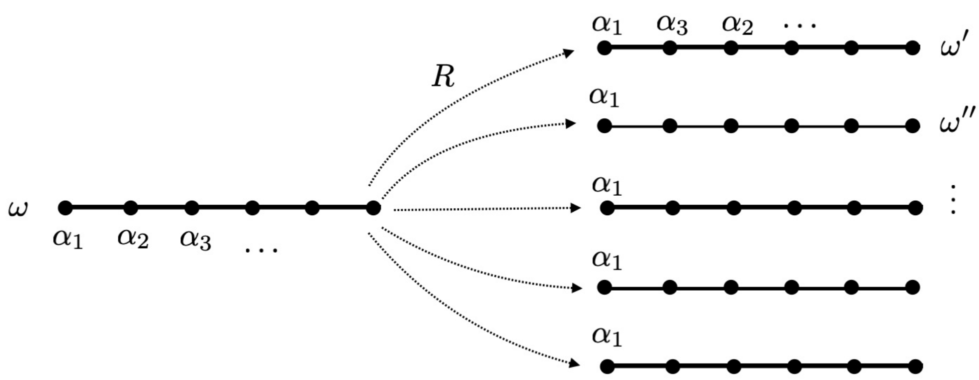


Figure 2.4: A pictorial representation of the accessible atoms from a given one ω in a Lewis frame.

It is easy to observe, also graphically, that, given a Lewis frame $\langle at(C(\mathbf{A})), R\rangle$, for each $\omega \in at(C(\mathbf{A}))$, $R[\omega]$ is a set of permutations beginning with the same element (by **Ser** and **Cen**). The question we address now is: what is the consequence of imposing **Sph** on $R[\omega]$? Specifically, what properties of $R[\omega]$ does **Sph** characterize? As we already proved above, **Cen** and **Ser** characterize over $R[\omega]$, the properties of non-emptyness and seriality, that is $R[\omega]$ must be non-empty, by **Ser**, and all the elements in $R[\omega]$ must begin with the same initial element as ω, by **Cen** (see Table 2.3.1). However, for now, we have only been able to describe the impact of **Sph** via the selection function tool f. We will now introduce a new construction that allows us to make the impact of **Sph** more transparent and show the property of $R[\omega]$ that **Sph** characterizes. We will make extensive use of graphical examples.

By Figure 2.3.2, it is easy to observe that the element of $R[\omega]$ can be arranged in a *matrix* as the next example shows.

Example of the matrix induced by $R[\omega]$

Consider a dual Kripke frame $\langle at(C(\mathbf{A})), R\rangle$ where:

- $at(C(\mathbf{A}))$ is the set of atoms of a BAC, $C(\mathbf{A})$, generated by an algebra $\mathbf{A}$ with 6 atoms $at(\mathbf{A}) = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6\}$, i.e. $|at(C(\mathbf{A}))| = 6!$;
- consider an element $\omega \in at(C(\mathbf{A}))$ such that $\omega[1] = \alpha_1$, i.e. the first element appearing in ω is α_1
- R is such that $R[\omega] = \{\langle\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6\rangle, \langle\alpha_1, \alpha_3, \alpha_2, \alpha_4, \alpha_6, \alpha_5\rangle, \langle\alpha_1, \alpha_4, \alpha_2, \alpha_3, \alpha_5, \alpha_6\rangle\}$

The $R[\omega]$ can be arranged into the following 3×6 matrix, where $3 = |R[\omega]|$ and $6 = |at(\mathbf{A})|$:

$$\begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_5 & \alpha_6 \\ \alpha_1 & \alpha_3 & \alpha_2 & \alpha_4 & \alpha_6 & \alpha_5 \\ \alpha_1 & \alpha_4 & \alpha_2 & \alpha_3 & \alpha_5 & \alpha_6 \end{pmatrix}$$

In general, let us use the following notation:

Notation 2.11. *Matrices and Lewis frames*

- *with use boldface upper-cases Latin latter to denote a matrix,* $\mathbf{M}$*,* $\mathbf{N}$*,....*
- *for a matrix* $\mathbf{M}$*, M in italic denotes the set of the elements appearing in* $\mathbf{M}$*;*
- *If* $\mathbf{M}$ *and* $\mathbf{N}$ *are matrices having the same number of rows,* $\mathbf{M} \cdot \mathbf{N}$ *denotes the matrix obtained by sewing together* $\mathbf{M}$ *an* $\mathbf{N}$ *along the row-side. For*

example:

$$\mathbf{M} = \begin{pmatrix} \alpha_1 & \alpha_2 \\ \alpha_1 & \alpha_3 \\ \alpha_1 & \alpha_4 \end{pmatrix} \cdot \mathbf{N} = \begin{pmatrix} \alpha_3 & \alpha_4 \\ \alpha_2 & \alpha_4 \\ \alpha_2 & \alpha_3 \end{pmatrix} = \mathbf{M} \cdot \mathbf{N} = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \\ \alpha_1 & \alpha_3 & \alpha_2 & \alpha_4 \\ \alpha_1 & \alpha_4 & \alpha_2 & \alpha_3 \end{pmatrix}$$

- *Given a matrix* $\mathbf{M}$, $\alpha_{i,j}$ *denotes the element of* $\mathbf{M}$ *appearing in the i-th row and in the j-th column*

We have now the ingredients to define more formally the definition of a matrix induced by the elements of a dual frame of a modal BAC:

Definition 2.14. *Consider a dual Kripke frame* $\langle at(\mathcal{C}(\mathbf{A})), R\rangle$ *of a modal BAC* $\langle \mathcal{C}(\mathbf{A}), \Box\rangle$ *and an element* $\omega \in at(\mathcal{C}(\mathbf{A}))$. *Assume* $|at(\mathbf{A})| = n$ *and* $|R[\omega]| = m$*:*

- $\mathbf{R}[\omega]$ *is the* $m \times n$ *matrix induced by* $R[\omega]$ *whose rows are the elements of* $R[\omega]$.

Furthermore, we use the following notation: for all $1 \leq i \leq k$ *and* $1 \leq j \leq n$*,*

- $\mathbf{r}_i$*, in boldface, denotes the i-th row of* $\mathbf{R}[\omega]$*; and* r_i*, in italic, denotes the set of elements appearing in* $\mathbf{r}_i$*;*

- $\mathbf{c}_j$*, in boldface, denotes the j-th column of* $\mathbf{R}[\omega]$*, and* c_j*, in italic, denotes the set of elements appearing in* $\mathbf{c}_j$

Now, we will introduce a construction that allows us to make some structural properties of $\mathbf{R}[\omega]$ more transparent:

Definition 2.15 (Partition of a Matrix)**.** *For a dual Kripke frame* $\langle at(\mathcal{C}(\mathbf{A})), R\rangle$ *of a modal BAC* $\langle \mathcal{C}(\mathbf{A}), \Box\rangle$ *and an element* $\omega \in at(\mathcal{C}(\mathbf{A}))$*, assume* $|at(\mathbf{A})| = n$ *and* $|R[\omega]| = m$ *and consider* $\mathbf{R}[\omega]$ *whose dimensions are* $m \times n$ *according to Definition 2.14.*

A partition of $\mathbf{R}[\omega]$ *is a set* Π *of sub-matrices of* $\mathbf{R}[\omega]$ *defined by the following inductive procedure:*

- ***Starting Point****. Let us start with* $\Pi = \emptyset$*. The following instructions define a procedure to append elements to* Π *step by step.*

- ***Basic Step****. Consider* $\mathbf{c}_1$ *in* $\mathbf{R}[\omega]$*, i.e. the first column in* $\mathbf{R}[\omega]$*; assume* $|c_1| = m_1$*, i.e.* m_1 *is the number of elements appearing in* $\mathbf{c}_1$.

  – *If* $m_1 = n$*, i.e.* $m_1 = |at(\mathbf{A})|$*, then append the whole* $\mathbf{R}[\omega]$ *to* Π *and stop the procedure. The final output will be* $\Pi = \{\mathbf{R}[\omega]\}$.

  - *if $m_1 < n$, let $\mathbf{C}_1$ denote the sub-matrix of $\mathbf{R}[\omega]$ whose columns are, in the order, $\mathbf{c_1}, \dots, \mathbf{c}_{m_1}$. Then append $\mathbf{C}_1$ to Π, obtaining $\Pi = \{\mathbf{C}_1\}$ and proceed to the next inductive step.*

- ***Inductive Step.*** *Consider the first column $\mathbf{c}_i$ in $\mathbf{R}[\omega]$ such that $\mathbf{c}_i$ does not already appear in any of elements of Π at this stage. Assume $|c_i| = m_i$, then we have different cases:*
  - *if $i + m_i - 1 < n$, append to Π the sub-matrix $\mathbf{C}_i$ whose columns are, in the order, $\mathbf{c}_i, \dots, \mathbf{c}_{i+m_i-1}$ and repeat this **Inductive Step**;*
  - *if $i + m_i - 1 \geq n$, append to Π the sub-matrix $\mathbf{C}_i$ whose columns are, in the order, $\mathbf{c}_i, \dots, \mathbf{c}_n$ and stop the procedure.*

  The final output will be a partition of the form $\Pi = \{\mathbf{C}_1, \dots, \mathbf{C}_k\}$

Clearly, the above procedure stops and outputs the desired outcomes since $\mathbf{A}$ is finite, as well as the set $R[\omega]$.
Furthermore, it is readily observable that, if the final output is $\Pi = \{\mathbf{C}_1, \dots, \mathbf{C}_k\}$, then $\mathbf{C}_1 \cdot \dots \cdot \mathbf{C}_k = \mathbf{R}[\omega]$. That is, zipping together the elements of Π results in $\mathbf{R}[\omega]$.

Roughly, a partition of a matrix $\mathbf{R}[\omega]$, according to the above definition, consists in dividing $\mathbf{R}[\omega]$ into sub-matrices such that each of them has the same number of rows as $\mathbf{R}[\omega]$ and as many columns as the number of elements appearing in its first column, except, possibly, for the last sub-matrix of the partition. Before introducing an example of this construction, let us define a special property of the partition of a matrix:

Definition 2.16. *For a dual Kripke frame $\langle at(C(\mathbf{A})), R\rangle$ of a modal BAC $\langle C(\mathbf{A}), \Box\rangle$ and an element $\omega \in at(C(\mathbf{A}))$, consider $\mathbf{R}[\omega]$ and its partition $\Pi = \{\mathbf{C}_1, \dots, \mathbf{C}_n\}$ according to Definition 2.15. We define the following property of $\mathbf{R}[\omega]$:*

$\mathbf{R}[\omega]$ *is **Spherical*** $\Leftrightarrow$ *for each $\mathbf{C}_i \in \Pi$, the elements appearing in the first column of $\mathbf{C}_i$ are exactly the same elements appearing in each row of $\mathbf{C}_i$.*
More formally, let $\mathbf{c}_i$ be the first row of $\mathbf{C}_i$: for all rows $\mathbf{r}$ of $\mathbf{C}_i$, $c_i = r$.

*if $\mathbf{R}[\omega]$ is **Spherical**, we may refer to Π as a spherical partition over $\mathbf{R}[\omega]$. Furthermore we define the same property for the whole frame $\langle at(C(\mathbf{A})), R\rangle$ as follows:*

$$\langle at(C(\mathbf{A})), R\rangle \textit{ is } \textbf{\textit{Spherical}} \quad \Leftrightarrow \quad \textit{for all } \omega \in at(C(\mathbf{A})),\ \mathbf{R}[\omega] \textit{ is } \textbf{\textit{Spherical}}$$

The following example will clarify the partition construction and the property of **Sphericality**. 2.3.2

Example of (non-)Spherical Partitions

Consider a Lewis frame $\langle at(C(\mathbf{A})), R\rangle$ such that:

- $at(C(\mathbf{A}))$ is the set of atoms of a BAC, $C(\mathbf{A})$, generated by an algebra $\mathbf{A}$ with 6 atoms $at(\mathbf{A}) = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6\}$, i.e. $|at(C(\mathbf{A}))| = 6!$;
- consider two elements $\omega, \omega' \in at(C(\mathbf{A}))$ such that $\omega[1] = \omega'[1] = \alpha_1$, i.e. the first element appearing in ω and ω' is α_1
- R is such that
  $R[\omega] = \{\langle\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6\rangle, \langle\alpha_1, \alpha_3, \alpha_2, \alpha_4, \alpha_6, \alpha_5\rangle, \langle\alpha_1, \alpha_4, \alpha_2, \alpha_3, \alpha_5, \alpha_6\rangle\}$
  $R[\omega'] = \{\langle\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6\rangle, \langle\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_6, \alpha_5\rangle, \langle\alpha_1, \alpha_5, \alpha_2, \alpha_3, \alpha_4, \alpha_6\rangle\}$

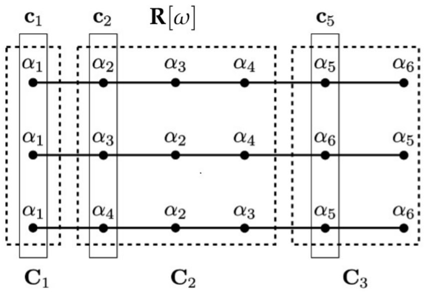
$R[\omega]$ can be arranged into the following 3×6 matrix, where $3 = |R[\omega]|$ and $6 = |at(\mathbf{A})|$:

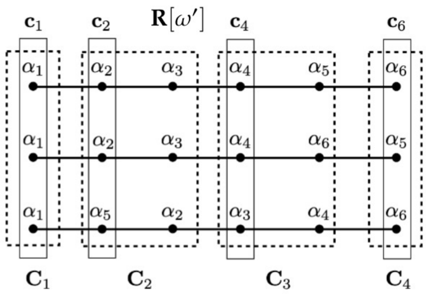
$$\begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_5 & \alpha_6 \\ \alpha_1 & \alpha_3 & \alpha_2 & \alpha_4 & \alpha_6 & \alpha_5 \\ \alpha_1 & \alpha_4 & \alpha_2 & \alpha_3 & \alpha_5 & \alpha_6 \end{pmatrix}$$

$R[\omega']$ can be arranged into the following 3×6 matrix, where $3 = |R[\omega]|$ and $6 = |at(\mathbf{A})|$:

$$\begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_5 & \alpha_6 \\ \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_6 & \alpha_5 \\ \alpha_1 & \alpha_5 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_6 \end{pmatrix}$$

Consider the partitions of $\mathbf{R}[\omega]$ and $\mathbf{R}[\omega']$ as in the following picture:





- $\Pi = \{\mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3\}$ is the partition of $\mathbf{R}[\omega]$, on the left, defined by following the procedure in Definition 2.15:
  - since $|c_1| = 1$, $\mathbf{C}_1 = \mathbf{c}_1$
  - since $|c_2| = 3$, $\mathbf{C}_2$ is made of all the three columns after $\mathbf{C}_1$, i.e. $\mathbf{C}_2 = \mathbf{c}_2 \cdot \mathbf{c}_3 \cdot \mathbf{c}_4$

- since $|c_5| = 2$, $\mathbf{C}_3$ is made of all the two columns after $\mathbf{C}_2$, i.e. $\mathbf{C}_3 = \mathbf{c}_5 \cdot \mathbf{c}_6$

- $\Pi' = \{\mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3, \mathbf{C}_4\}$ is the partition of $\mathbf{R}[\omega']$, on the left, defined by following the procedure in Definition 2.15:
  - since $|c_1| = 1$, $\mathbf{C}_1 = \mathbf{c}_1$
  - since $|c_2| = 2$, $\mathbf{C}_2$ is made of all the two columns after $\mathbf{C}_1$, i.e. $\mathbf{C}_2 = \mathbf{c}_2 \cdot \mathbf{c}_3$
  - since $|c_3| = 2$, $\mathbf{C}_3$ is made of all the two columns after $\mathbf{C}_2$, i.e. $\mathbf{C}_3 = \mathbf{c}_4 \cdot \mathbf{c}_5$
  - observe that $|c_6| = 2$, so $2+5 > 6$, namely the number of elements in c_6, which is 2, exceeds the number of remaining columns after $\mathbf{c}_5$, which is just $\mathbf{c}_6$. Hence, the last sub-matrix is made of all the remaining columns anyway, which is just $\mathbf{c}_6$, i.e. $\mathbf{C}_4 = \mathbf{c}_6$.

- Π is **Spherical** since for all its sub-matrices $\mathbf{C}_i$, the elements appearing in the first column of $\mathbf{C}_i$ are exactly those appearing in each of the rows of $\mathbf{C}_i$:
  - $\mathbf{C}_1 = \mathbf{c}_1$ and $c_1 = \{\alpha_1\}$, hence the relevant property clearly holds for $\mathbf{C}_1$
  - the first column of $\mathbf{C}_2$ is $\mathbf{c}_2$ and $c_2 = \{\alpha_2, \alpha_3, \alpha_4\}$ and these elements appear in each of the rows of $\mathbf{C}_2$, as the picture shows;
  - the first column of $\mathbf{C}_3$ is $\mathbf{c}_5$ and $c_5 = \{\alpha_5, \alpha_6\}$ and these elements appear in each of the rows of $\mathbf{C}_3$, as the picture shows.

- Π' is not **Spherical** since there is a sub-matrix $\mathbf{C}_2$ where the elements appearing in its first columns are not exactly those appearing in each of the rows of $\mathbf{C}_i$. Specifically:
  - the first column of $\mathbf{C}_2$ is $\mathbf{c}_2$ and $c_2 = \{\alpha_2, \alpha_5\}$; however the middle row of $\mathbf{C}_2$ only contains the elements in $\{\alpha_2, \alpha_3\}$.

From Definition 2.15 and Lemma 2.7, the following remark readily follows:

Remark 2.2. *For a dual Kripke frame $\langle at(C(\mathbf{A})), R\rangle$ of a modal BAC, and for $\omega \in at(C(\mathbf{A}))$, consider $\mathbf{R}[\omega]$ and assume it is* **Spherical**. *The partition Π of $\mathbf{R}[\omega]$ is such that for all $\mathbf{C}_i$, all the rows of $\mathbf{C}_i$ are permutations of the elements of the first column of $\mathbf{C}_i$.*

We are approaching the more transparent characterization of the property induced by the condition $(L3)$. First, we require the following technical lemma:

Lemma 2.18. *Let $\langle at(C(\mathbf{A})), R\rangle$ be a dual Kripke frame of a modal BAC; consider a $\omega \in at(C(\mathbf{A}))$, the matrix $\mathbf{R}[\omega]$ and the partition Π of $\mathbf{R}[\omega]$. Assume that $\mathbf{R}[\omega]$ is not **Spherical**; Since Π is finite, let $\mathbf{C}_l$ denote the first sub-matrix of Π witnessing the failure of **Sphericality**, and let $\mathbf{c}_l$ denote the first column of $\mathbf{C}_l$. So, there exists a row $\mathbf{r}$ in $\mathbf{C}_l$ such that $c_l \neq r$.*

Thus, the following hold:

1. *$|c_l| \geq 2$;*
2. *There is a $\alpha_{i,j} \in r$ such that $\alpha_{i,j} \notin c_l$;*
3. *For all $\alpha_{i,j} \in r$ such that $\alpha_{i,j} \notin c_l$, there is $\alpha_{x,y} \in c_l$ such that $y \neq j$ and $\alpha_{x,y}$ does not appear in the j-th row of $\mathbf{R}[\omega]$ earlier than $\alpha_{i,j}$.*

Proof. The proof is included in the Appendix B.2. □

We can now prove the following characterization:

Theorem 2.5. *For a dual Kripke frame $\langle at(C(\mathbf{A})), R\rangle$ of a modal BAC, the following holds:*

$$\langle at(C(\mathbf{A})), R\rangle \text{ satisfies } \mathbf{Sph} \text{ from Table 2.3.1} \quad \Leftrightarrow \quad \text{for all } \omega \in at(C(\mathbf{A})),\ \mathbf{R}[\omega] \text{ is } \mathbf{Spherical}$$

Proof.

(⇒) By contraposition, assume that there is $\omega \in at(C(\mathbf{A}))$ for which $\mathbf{R}[\omega]$ does not admit a spherical partition. Furthermore, assume that $\mathbf{C}_l$ is the first submatrix of $\mathbf{R}[\omega]$ for which $c_l \neq r$ for a row $\mathbf{r}$ of $\mathbf{C}_l$. Then, by Lemma 2.18-1-2 we know that c_l contains (at least) two different elements $\alpha_s \neq \alpha_p$ and there exists $\alpha_{i,j}$ such that $\alpha_{i,j} \in r$ but it does not appear in c_l. Then, since $\mathbf{c}_l$ is the first column of $\mathbf{R}[\omega]$ that witness the failure of **Sphericality**, $\alpha_{i,j}$ does not belong to any other column that precedes $\mathbf{c}_l$ in the order of columns of $\mathbf{R}[\omega]$. Moreover, by Lemma 2.18-3, there exists an $\alpha_{x,y} \in c_l$ such that $y \neq j$ and $\alpha_{x,y}$ does not appear in the jth row of $\mathbf{R}[\omega]$, before the ith column, i.e., $\alpha_{x,y}$ does not appear in the jth row before $\alpha_{i,j}$.

Then, let $a = (\alpha_{x,y} \vee \alpha_{i,j})$ and $b = \bigvee(c_l \setminus \{\alpha_{x,y}\})$. Notice that $[a] = \{\alpha_{x,y}, \alpha_{i,j}\}$ and $[b] = c_l \setminus \{\alpha_{x,y}\}$. Let f be the selection function as in Definition 2.13. Notice that,

$$f(a \vee b, \omega) = \{\omega'[i] \mid \omega R \omega', \omega'[i] \leq a \vee b, \text{ and for all } j < i, \omega'[j] \nleq a \vee b\} = c_l$$

Clearly $f(a \vee b, \omega) \not\subseteq [a]$ since $\alpha_{i,j} \notin c_l$. Also $f(a \vee b, \omega) \not\subseteq [b]$ since $\alpha_{x,y} \notin [b]$. Moreover, $\alpha_{i,j} \in f(a, \omega)$ since $\alpha_{x,y}$ does not appear before on the same j-th row as $\alpha_{i,j}$, and $\alpha_{i,j}$ does not appear anywhere in $\mathbf{R}[\omega]$ before $\mathbf{c}_l$. Hence, $f(a, \omega) \cup f(b, \omega) \not\subseteq f(a \vee b, \omega)$.

($\Leftarrow$) Assume that F is spherical, and hence for all $\omega \in at(C(\mathbf{A}))$, $\Pi = \{\mathbf{C}_1, \mathbf{C}_2, \ldots, \mathbf{C}_t\}$ is a spherical partition of $\mathbf{R}[\omega]$. We now prove that for all $\omega \in at(C(\mathbf{A}))$ and $a, b \in A$, $f(a \vee b, \omega) \not\subseteq [a]$ and $f(a \vee b, \omega) \not\subseteq [b]$ implies that $f(a, \omega) \cup f(b, \omega) \subseteq f(a \vee b, \omega)$. Notice that the conditions $f(a \vee b, \omega) \not\subseteq [a]$ and $f(a \vee b, \omega) \not\subseteq [b]$ are equivalent to the existence of $\alpha, \beta \in at(\mathbf{A})$ such that $\alpha, \beta \in f(a \vee b, \omega)$ and $\alpha \in [a]$ and $\alpha \notin [b]$, and $\beta \in [b]$ but $\beta \notin [a]$, i.e., $\alpha \leq a \wedge \neg b$ and $\beta \leq b \wedge \neg a$. We distinguish two cases:

(1) α, β are in the same $\mathbf{C}_l \in \Pi$, more precisely, $\alpha, \beta \in C_l$. This means that α and β appears in each row of $\mathbf{C}_l$ by Remark 2.2. Now, $\alpha, \beta \in f(a \vee b, \omega)$, implies $f(a \vee b, \omega) \subseteq C_l$. Indeed, by way of contradiction, assume there is $\gamma \in f(a \vee b, \omega)$ such that $\gamma \notin C_l$. If $\gamma \in C_z$ for some $z > l$, by definition of the selection function f, α or β would appear before γ on the same row, contradicting the assumption that $\gamma \in f(a \vee b, \omega)$. Thus, assume that $\gamma \in C_z$ for some $z < l$. In this case, by Remark 2.2 again, γ appears in each row of $\mathbf{C}_z$, and so, it must appears before α (β), in the same rows as α (β). This implies that $\alpha \neq \gamma$ and $\beta \neq \gamma$ and $\alpha, \beta \notin f(a \vee b, \omega)$, contradicting our assumption. So, $f(a \vee b, \omega) \subseteq C_l$. By an analogous reasoning and using the fact that $\alpha \in [a]$ and $\alpha \notin [b]$, and $\beta \in [b]$ but $\beta \notin [a]$, we can show that $f(a, \omega) \subseteq C_l$ and $f(b, \omega) \subseteq C_l$.

Let $\mathbf{c}_l$ be the first column of $\mathbf{C}_l$ so that $f(a, \omega) \subseteq c_l = C_l$, $f(b, \omega) \subseteq c_l = C_l$, and $f(a \vee b, \omega) \subseteq c_l = C_l$. Hence, $f(a, \omega), f(b, \omega), f(a \vee b, \omega)$ are all subsets of c_l. Therefore, it cannot exists $\gamma \in f(a \vee b, \omega)$ but $\gamma \notin f(a, \omega)$ and $\gamma \notin f(b, \omega)$ because if $\gamma \leq a \vee b$, then $\gamma \leq a$, $\gamma \leq b$ and $\gamma \in c_l$. And this implies $f(a, \omega) \cup f(b, \omega) \subseteq f(a \vee b, \omega)$.

(2) α, β are in two different submatrices in Π, i.e. for some $\mathbf{C}_l \neq \mathbf{C}_p \in \Pi$, $\alpha \in C_l$, and $\beta \in C_p$. Thus, $\beta \notin C_l$. By contradiction, we show that this case cannot hold. Indeed, and without loss of generality, assume $l < p$, then, by definition of Π, α appears in each row of $\mathbf{C}_l$. By reasoning analogously to the case (1) above, we can show that $f(a \vee b, \omega) \subseteq C_l$. This leads to a contradiction with the fact that $\beta \in f(a \vee b, \omega)$ but $\beta \notin C_l$.

□

The above results show the the property of **Sph** in Table 2.3.1 is equivalent to a certain peculiar property of the accessibility relation. In particular, from Theorem 2.5 and Lemma 2.16 we derive the following corollary:

Corollary 2.5. *Consider a modal BAC* $\langle C(\mathbf{A}), \Box\rangle$ *and its dual Kripke frame* $\langle at(C(\mathbf{A})), R\rangle$*. The following are equivalent:*

- $\langle C(\mathbf{A}), \Box\rangle$ *satisfies* $(L3)$ *as in Definition 2.12;*
- $\langle at(C(\mathbf{A})), R\rangle$ *satisfies* **Sph** *as in Table 2.3.1;*
- $\langle at(C(\mathbf{A})), R\rangle$ *is* **Spherical** *as in Definition 2.16.*

Therefore, we can rewrite the table describing dual modal frames as follows:

	Lewis Frames: $\langle at(C(\mathbf{A})), R\rangle$	
	Conditions on R	Conditions on Lewis algebra: $\langle C(\mathbf{A}), R\rangle$
(Ser) Seriality	for all $\omega \in at(C(\mathbf{A}))$, there is a ω' such that $\omega R \omega'$	$\Box(a \mid \top) = (a \mid \top)$
(Cen) Centering	for all $\omega, \omega' \in at(C(\mathbf{A}))$, if $\omega R \omega'$ then $\omega[1] = \omega[1]$	
(Sph) Sphericality	for all $a, b \in A$, for all $\omega \in at(C(\mathbf{A}))$, either $f(a \vee b, \omega) \subseteq [a]$ or $f(a \vee b, \omega) \subseteq [b]$ or $f(a \vee b, \omega) = f(a, \omega) \cup f(b, \omega)$ for all $\omega \in at(C(\mathbf{A}))$, $\mathbf{R}[\omega]$ is **Spherical** as in Definition 2.16	$(L3)$ in Definition 2.12

Table 2.3: The table describes the properties over a dual Kripke frame of a modal BAC specifying the implicit version (employing the selection function f) and the transparent version (employing the accessibility relation) of the property induced by $(L3)$ in Definition 2.12

We have introduced all the necessary structural properties of Lewis algebras to conduct our investigation of conditional logics inside the modal BACs framework. As we already anticipated, the next results will show how we can interpret a Lewis counterfactual $a \Box\!\!\rightarrow b$ into a modal conditional object of the form $\Box(a \mid b)$.

2.4 Logics and Modal BACs

In this section we will show how to construct interpretations for a new modal extension of the language $\mathcal{L}_{LBC}$ and examine the resulting logics. It is essential to recall the definition of LBC-valuation, (see Definition 2.9) an the properties of the modal BAC induced by $\mathbf{L}$, i.e. the Lindenbaum algebra of CPL over the language $\mathcal{L}$ (see Lemma 2.11 and Corollary 2.3).

2.4.1 Syntax and Semantics

We introduce a new language resulting from a modal expansion of $\mathcal{L}_{LBC}$:

Definition 2.17 ($LBC_{\Box}$ language). *$\mathcal{L}_{LBC_{\Box}}$ is a language obtained by expanding $\mathcal{L}_{LBC}$ with the modal operator $\Box$. Formulas of $\mathcal{L}_{LBC}$ are defined as follows:*

- *if Φ is a formula of $\mathcal{L}_{LBC}$, then Φ is a formula of $\mathcal{L}_{LBC_{\Box}}$*
- *if Φ, Ψ are formulas of $\mathcal{L}_{LBC_{\Box}}$, then so are $\Box\Phi$, $\Phi \wedge \Psi$, $\Phi \vee \Psi$, and $\neg\Phi$;*
- *nothing else is a formula of $\mathcal{L}_{LBC_{\Box}}$.*

$For_{\mathcal{L}_{LBC}}$ denotes the set of formulas of $\mathcal{L}_{LBC}$. Moreover, we can defined the $\Diamond$ operator as customary: $\Diamond\Phi := \neg\Box\neg\Phi$

Furthemore, we can define the following translation from formulas of $\mathcal{L}^{\upharpoonright}_{\Box\!\!\rightarrow}$ into formulas of $\mathcal{L}_{LBC_{\Box}}$.

Definition 2.18. *The function $\tau : For_{\mathcal{L}^{\upharpoonright}_{\Box\!\!\rightarrow}} \to For_{\mathcal{L}_{LBC_{\Box}}}$ is a translation of the language $\mathcal{L}^{\upharpoonright}_{\Box\!\!\rightarrow}$ into the language $\mathcal{L}_{LBC_{\Box}}$ is inductively defined as follows:*

- *for $p \in Var$, $\tau(p) = (p \mid \top)$*
- *for $\varphi, \psi \in For_{\mathcal{L}}$ with $\nvdash_{CPL} \neg\varphi$, $\tau(\varphi \Box\!\!\rightarrow \psi) = \Box(\psi \mid \varphi)$*
- *for $\varphi, \psi \in For_{\mathcal{L}^{\upharpoonright}_{\Box\!\!\rightarrow}}$, $\tau(\varphi \wedge \psi) = \tau(\varphi) \wedge \tau(\psi)$*
- *for $\varphi, \psi \in For_{\mathcal{L}^{\upharpoonright}_{\Box\!\!\rightarrow}}$, $\tau(\varphi \vee \psi) = \tau(\varphi) \vee \tau(\psi)$*
- *for $\varphi \in For_{\mathcal{L}^{\upharpoonright}_{\Box\!\!\rightarrow}}$, $\tau(\neg\varphi) = \neg\tau(\varphi)$*

The above translation establishes that classical formulas in $\mathcal{L}^{\upharpoonright}_{\Box\!\!\rightarrow}$ can be interpreted as conditional formulas in $\mathcal{L}_{LBC_{\Box}}$ having $\top$ as the "antecedent", while counterfactual formulas of the form $\varphi \Box\!\!\rightarrow \psi$ can be interpreted as modal conditionals of the form $\Box(\psi \mid \varphi)$, and Boolean combinations of classical and counterfactuals formulas can be interpreted as Boolean combinations of the corresponding translated formulas. Recall that $\mathcal{L}^{\upharpoonright}_{\Box\!\!\rightarrow}$ doesn't allow for nested occurrences of the counterfactual connective, hence τ is well defined.

Example of a translation τ

- $\tau((p \vee q) \Box\!\!\rightarrow r) = \Box(r \mid p \vee q)$
- $\tau(p \vee \neg(q \vee r)) = (p \mid \top) \vee (\neg(q \mid \top) \vee (r \mid \top))$
- $\tau((p \Box\!\!\rightarrow q) \wedge (r \vee q)) = \Box(q \mid p) \wedge ((r \mid \top) \vee (q \mid \top))$
- $\tau(\neg(\neg(p \Box\!\!\rightarrow q) \vee (p \wedge \neg r))) = \neg(\neg\Box(q \mid p) \vee ((p \mid \top) \wedge (\neg r \mid \top)))$

The semantics for $\mathcal{L}_{LBC_\Box}$ is a natural modal extension of the semantics for $\mathcal{L}_{LBC}$. First, we need to introduce a special kind of Lewis algebra into which $\mathcal{L}_{LBC_\Box}$ can be interpreted.

Definition 2.19. *A canonical Lewis algebra is a Lewis algebra of the form* $\langle C(\mathbf{L}), \Box \rangle$ *where* $\mathbf{L}$ *is the Lindenbaum algebra of CPL over* $\mathcal{L}$ *(recall Corollary 2.3).*

Observe that there might be different canonical Lewis algebras depending on the different operators $\Box$ we one can define over $C(\mathbf{L})$. It is clear to see that also for a canonical Lewis algebra, $\langle C(\mathbf{L}), \Box \rangle$, the correspondence between permutations of classical valuations and atoms of $\langle C(\mathbf{L}), \Box \rangle$ holds:

Corollary 2.6. *Consider a canonical Lewis algebra,* $\langle C(\mathbf{L}), \Box \rangle$*. There is a bijection* $\Omega^* : Perm(Val_{CPL}) \rightarrowtail\!\!\!\!\twoheadrightarrow at(C(\mathbf{L}))$ *between atoms of* $\langle C(\mathbf{L}), \Box \rangle$ *and permutations of classical valuations such that*

$$\Omega^*(\langle v_1, \ldots, v_n \rangle) = \Omega(\langle v_1^*, \ldots, v_n^* \rangle)$$

where * *is the same bijection from Lemma 2.11 and* Ω *is the same bijection from Lemma 2.7.*

Thus, for simplicity, we will thereby identify atoms of $C(\mathbf{L})$ with permutations of classical valuations of $\mathcal{L}$

Proof. Straightforward from Corollary 2.3. □

Canonical Lewis algebras can be employed to interpret $\mathcal{L}_{LBC_\Box}$ in the following way:

Definition 2.20. *For a canonical Lewis algebra* $\langle C(\mathbf{L}), \Box \rangle$*, Let* $[]_{LBC_\Box} : For_{\mathcal{L}_{LBC}} \to \langle C(\mathbf{L}), \Box \rangle$ *be an interpretation from formulas of* $\mathcal{L}_{LBC_\Box}$ *to* $\langle C(\mathbf{L}), \Box \rangle$ *inductively defined as follows:*

- *if* $\Phi \in For_{\mathcal{L}_{LBC}}$*, then* $[\Phi]_{LBC_\Box} = [\Phi]_{LBC}$ *according to Definition 2.9;*
- $[\Box\Phi]_{LBC_\Box} = \Box([\Phi]_{LBC_\Box})$
- $[\Phi \wedge \Psi]_{LBC_\Box} = [\Phi]_{LBC_\Box} \sqcap [\Psi]_{LBC_\Box}$
- $[\Phi \vee \Psi]_{LBC_\Box} = [\Phi]_{LBC_\Box} \sqcup [\Psi]_{LBC_\Box}$
- $[\neg\Phi]_{LBC} = \sim [\Phi]_{LBC_\Box}$

Namely, $[]_{LBC_\Box}$ is a natural modal extension of the interpretation $[]_{LBC}$ for the language $\mathcal{L}_{LBC_\Box}$. We have now all the ingredients to define $LBC_\Box$-valuations:

Definition 2.21 ($LBC_{\Box}$-valuation). *Consider a canonical Lewis algebra, $\langle C(\mathbf{L}), \Box \rangle$, and the interpretation $[]_{LBC_{\Box}} : For_{\mathcal{L}_{LBC_{\Box}}} \to \langle C(\mathbf{L}), \Box \rangle$ as in Definition 2.20.*

A $LBC_{\Box}$-valuation is a Kripke model of the form $\langle Perm(Val_{CPL}), R, \vDash \rangle$ where:

- *$\langle Perm(Val_{CPL}), R \rangle$ is (up to isomorphism) the dual Lewis frame $\langle at(C(\mathbf{L})), R \rangle$ of $\langle C(\mathbf{L}), \Box \rangle$ given the bijection between $Perm(Val_{LBC})$ and $at(C(\mathbf{L}))$ by Corollary 2.6. (see also Definition 2.9).*
- *$\vDash \subseteq Perm(Val_{CPL}) \times For_{\mathcal{L}_{LBC_{\Box}}}$ is a satisfaction relation defined as in Definition 2.9: for all $e \in Perm(Val_{CPL})$, for all $\Phi \in For_{\mathcal{L}_{LBC_{\Box}}}$,*

$$e \vDash \Phi \Leftrightarrow \Omega^*(e) \sqsubseteq [\Phi]_{LBC_{\Box}}$$

As for the non modal case, the following semantic clauses are readily provable:

Lemma 2.19. *For any $LBC_{\Box}$-valuation $\langle Perm(Val_{CPL}), R, \vDash \rangle$, for all $e \in Perm(Val_{CPL})$, for all formulas $\Phi, \Psi \in For_{\mathcal{L}_{LBC_{\Box}}}$, the following holds:*

$$\begin{array}{lcl} e \vDash (\varphi \mid \psi) & \Leftrightarrow & \text{there is a } i \text{ such that } e[i](\varphi \wedge \psi) = 1 \text{ and for all } j < i, e[j](\psi) = 0 \\ e \vDash \Phi \wedge \Psi & \Leftrightarrow & e \vDash \Phi \text{ and } e \vDash \Psi \\ e \vDash \Phi \vee \Psi & \Leftrightarrow & e \vDash \Phi \text{ or } e \vDash \Psi \\ e \vDash \neg \Phi & \Leftrightarrow & e \nvDash \Phi \\ e \vDash \Box \Phi & \Leftrightarrow & \text{for all } e' \text{ such that } eRe', e' \vDash \Phi \end{array}$$

We adopt the following notation for logical consequence among formulas in $\mathcal{L}_{LBC_{\Box}}$:

Notation 2.12.

For $\Xi \cup \{\Phi\} \subseteq For_{\mathcal{L}_{LBC_{\Box}}}$,

$$\Xi \models_{LBC_{\Box}} \Phi \quad \Leftrightarrow \quad \begin{array}{l} \text{for all } LBC_{\Box}\text{-valuations } \langle Perm(Val_{CPL}), R, \vDash \rangle, \\ \text{for all } e \in Perm(Val_{CPL}), \text{ if } e \vDash \Psi \text{ for all } \Psi \in \Xi, \text{ then } e \vDash \Phi \end{array}$$

It is straightforward to see that $\models_{LBC_{\Box}}$ is a modal extension of $\models_{LBC}$. Indeed, $For_{\mathcal{L}_{LBC_{\Box}}}$ contains formulas of $\mathcal{L}_{LBC}$ too, and the logical behavior of (non-modal) LBC-formulas is preserved by the modal $\models_{LBC_{\Box}}$ consequence. In particular, it is straightforward to see that the logic **LBC**, which coincides with $\models_{LBC}$ is sound with respect to $\models_{LBC_{\Box}}$

Remark 2.3. *The logic $\models_{LBC}$ is sound with respect to $\models_{LBC_{\Box}}$ consequence: for $\Xi \cup \{\Phi\} \subseteq For_{\mathcal{L}_{LBC}}$,*

$$\Xi \vdash_{\mathbf{LBC}} \Phi \Rightarrow \Xi \models_{LBC_{\Box}} \Phi$$

In what follows, we will show how to interpret Lewis counterfactuals and the logic **VC**$^+$ by employing the $\mathcal{L}_{LBC_\Box}$ semantics.

2.4.2 From Spheres to $LBC_\Box$-valuations

In this section, we introduce a procedure to define a $LBC_\Box$-valuation out of a restricted spherical Model:

Definition 2.22. *Consider a (finite) restricted canonical spherical Lewisian model* $\mathcal{M}^{\mathfrak{C}} = \langle Val_{CPL}, \mathcal{S}^{\mathfrak{C}}, \vDash^{\mathfrak{C}} \rangle$ *satisfying* **Centering** *and assume* $c([w])_{\approx} = w$. *Since* Val_{CPL} *is finite, we can index the elements in* $\mathcal{S}^{\mathfrak{C}}(v)$ *from the smallest to the greatest, namely* $\mathcal{S}^{\mathfrak{C}}(v) = \{S_1^v, S_2^v, \ldots, S_n^v\}$ *such that* $S_1^v \subseteq S_2^v \subseteq \cdots \subseteq S_n^v$. *Now, define the set of radius of* v, $Radius_v$, *as follows:*

$$Rad_v = Perm(S_1^v) \times Perm(S_2^v \setminus S_1^v) \times \cdots \times Perm(S_n^v \setminus S_{n-1}^v)$$

For $r \in Rad_x$, $r[i]$ *denotes the i-th valuation appearing in* r, *then*

$$Radius_v = \{e \in Perm(Val_{CPL}) \mid \text{for some } r \in Rad_v, \text{for all } 1 \leq i \leq |Val_{CPL}|, e[i] = r[i]\}$$

However, r *can be identified with an element of* $Perm(Val_{CPL})$ *via the following bijective map* $F : Radius_x \rightarrowtail\!\!\!\twoheadrightarrow Perm(Val_{CPL})$: *for all* $r \in Radius_x$, *for all* $e \in Perm(Val_{CPL})$,

$$F(r) = e \Leftrightarrow \text{for all } 1 \leq i \leq |Val_{CPL}|, r[i] = e[i]$$

Now, define the $LBC_\Box$*-valuation induced by* $\mathcal{M}^{\mathfrak{C}}$ *as the tuple* $\langle Perm(Val_{CPL}), R, \vDash \rangle$ *where* R *is defined as follows:*

- *for all* $e \in Perm(Val_{CPL})$, $R[e] = \{r \mid r \in Radius_x \text{ and } x = e[1]\}$
- $\vDash \subseteq Perm(Val_{CPL}) \times For_{\mathcal{L}_{LBC_\Box}}$ *is defined according to Definition 2.21 and the clauses in Lemma 2.19.*

The above definition seems very complicated but it is based upon a very intuitive construction. First, notice that Each $r \in Rad_x$ is a tuple of permutations. F and Rad_x serve as "notational tools": consider for $e \in Perm(Val_{CPL})$,

$$F(r) = e \Leftrightarrow \text{for all } 1 \leq i \leq |Val_{CPL}|, r[i] = e[i]$$

Notice that r and e, with e being such that for all $1 \leq i \leq |Val_{CPL}|$ $e[i] = r[i]$, are essentially the same permutation of classical valuations. Hence, passing from Rad_v to $Radius_v$ is just a matter of notation: we transforms tuples in Rad_v into permutations in $Radius_v$. Furthermore, in the $LBC_\Box$-valuation induced by $\mathcal{M}^{\mathfrak{C}}$, $R[e]$ is the set of all permutations of classical valuations in $Radius_x$, whose firstly-appearing element is exactly the same as e's. In addition, observe that, by Jónsson-Tarski duality, the induced valuation $\langle Perm(Val_{CPL}), R, \vDash\rangle$ is (up to isomorphism) the dual Kripke frame induced by the canonical lewis algebra $\langle C(\mathbf{L}), \Box\rangle$ where $\Box X = \{e \in Perm(Val_{CPL}) \mid R[e] \subseteq X\}$.

Lemma 2.20. *The $LBC_\Box$-valuation $\langle Perm(Val_{CPL}), R, \vDash\rangle$ induced by a restricted canonical spherical model $\mathcal{M}^{\mathfrak{C}} = \langle Val_{CPl}, \mathcal{S}^{\mathfrak{C}}, \vDash^{\mathfrak{C}}\rangle$ satisfying* **Centering** *is indeed a $LBC_\Box$-valuation.*

Proof. The proof is included in Appendix B.3 □

The next example will clarify the above construction:

Example of a $LBC_\Box$-valuation induced by a restricted spherical model

Assume $Var = \{p, q\}$, thus $Val_{CPL} = \{v_1, v_2, v_3, v_4\}$ such that:

- $v_1(p) = 1$ and $v_1(q) = 1$;
- $v_2(p) = 1$ and $v_2(q) = 0$;
- $v_3(p) = 0$ and $v_3(q) = 1$;
- $v_4(p) = 0$ and $v_4(q) = 0$;

Consider the restricted canonical spherical model $\mathcal{M}^{\mathfrak{C}} = \langle Val_{CPL}, \mathcal{S}^{\mathfrak{C}}, \vDash^{\mathfrak{C}}\rangle$ satisfying **Centering** where:

- $Val_{CPL} = \{v_1, v_2, v_3, v_4\}$
- $\mathcal{S}^{\mathfrak{C}}(v_1) = \{ \underbrace{\{v_1\}}_{S_1^{v_1}}, \underbrace{\{v_1, v_2, v_3\}}_{S_2^{v_1}}, \underbrace{\{v_1, v_2, v_3, v_4\}}_{S_3^{v_1}}\}$
- $\mathcal{S}^{\mathfrak{C}}(v_v) = \{ \underbrace{\{v_2\}}_{S_1^{v_2}}, \underbrace{\{v_1, v_2, v_3\}}_{S_2^{v_2}}, \underbrace{\{v_1, v_2, v_3, v_4\}}_{S_3^{v_2}}\}$
- $\mathcal{S}^{\mathfrak{C}}(v_3) = \{ \underbrace{\{v_3\}}_{S_1^{v_3}}, \underbrace{\{v_1, v_2, v_3\}}_{S_2^{v_3}}, \underbrace{\{v_1, v_2, v_3, v_4\}}_{S_3^{v_3}}\}$
- $\mathcal{S}^{\mathfrak{C}}(v_4) = \{ \underbrace{\{v_4\}}_{S_1^{v_4}}, \underbrace{\{v_4, v_2, v_3\}}_{S_2^{v_4}}, \underbrace{\{v_1, v_2, v_3, v_4\}}_{S_3^{v_4}}\}$

The above model can be depicted as:

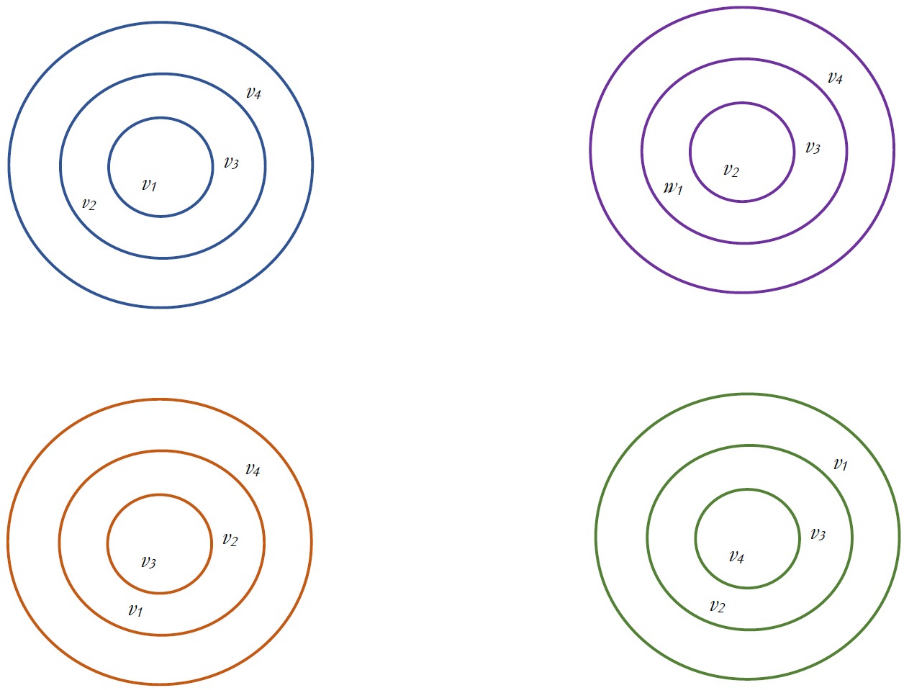


Figure 2.5: Example of a restricted spherical model.

Moreover, we have that:

- $Radius_{v_1} = \{\langle \underbrace{v_1}_{S_1^{v_1}}, \underbrace{v_2, v_3}_{S_2^{v_1} \backslash S_1^{v_1}}, \underbrace{v_4}_{S_3^{v_1} \backslash S_2^{v_1}} \rangle, \langle \underbrace{v_1}_{S_1^{v_1}}, \underbrace{v_3, v_2}_{S_2^{v_1} \backslash S_1^{v_1}}, \underbrace{v_4}_{S_3^{v_1} \backslash S_2^{v_1}} \rangle\}$
- $Radius_{v_2} = \{\langle \underbrace{v_2}_{S_1^{v_2}}, \underbrace{v_1, v_3}_{S_2^{v_2} \backslash S_1^{v_2}}, \underbrace{v_4}_{S_3^{v_2} \backslash S_2^{v_2}} \rangle, \langle \underbrace{v_2}_{S_1^{v_2}}, \underbrace{v_3, v_1}_{S_2^{v_2} \backslash S_1^{v_2}}, \underbrace{v_4}_{S_3^{v_2} \backslash S_2^{v_2}} \rangle\}$
- $Radius_{v_3} = \{\langle \underbrace{v_3}_{S_1^{v_3}}, \underbrace{v_1, v_2}_{S_2^{v_2} \backslash S_1^{v_3}}, \underbrace{v_4}_{S_3^{v_3} \backslash S_2^{v_3}} \rangle, \langle \underbrace{v_3}_{S_1^{v_3}}, \underbrace{v_2, v_1}_{S_2^{v_3} \backslash S_1^{v_3}}, \underbrace{v_4}_{S_3^{v_3} \backslash S_2^{v_3}} \rangle\}$
- $Radius_{v_4} = \{\langle \underbrace{v_4}_{S_1^{v_4}}, \underbrace{v_2, v_3}_{S_2^{v_3} \backslash S_1^{v_3}}, \underbrace{v_1}_{S_3^{v_4} \backslash S_2^{v_4}} \rangle, \langle \underbrace{v_4}_{S_1^{v_4}}, \underbrace{v_3, v_2}_{S_2^{v_4} \backslash S_1^{v_4}}, \underbrace{v_4}_{S_3^{v_4} \backslash S_2^{v_4}} \rangle\}$

As a result, the induced $LBC_{\Box}$-interpretation will be $\langle Perm(Val_{CPL}), R, \vDash \rangle$ where R is such that:

- for all $e \in Perm(Val_{CPL})$ such that $e[1] = v_1$, $R[e] = Radius_{v_1}$.

- for all $e \in Perm(Val_{CPL})$ such that $e[1] = v_2$, $R[e] = Radius_{v_2}$.
- for all $e \in Perm(Val_{CPL})$ such that $e[1] = v_3$, $R[e] = Radius_{v_3}$.
- for all $e \in Perm(Val_{CPL})$ such that $e[1] = v_4$, $R[e] = Radius_{v_4}$.

in a picture:

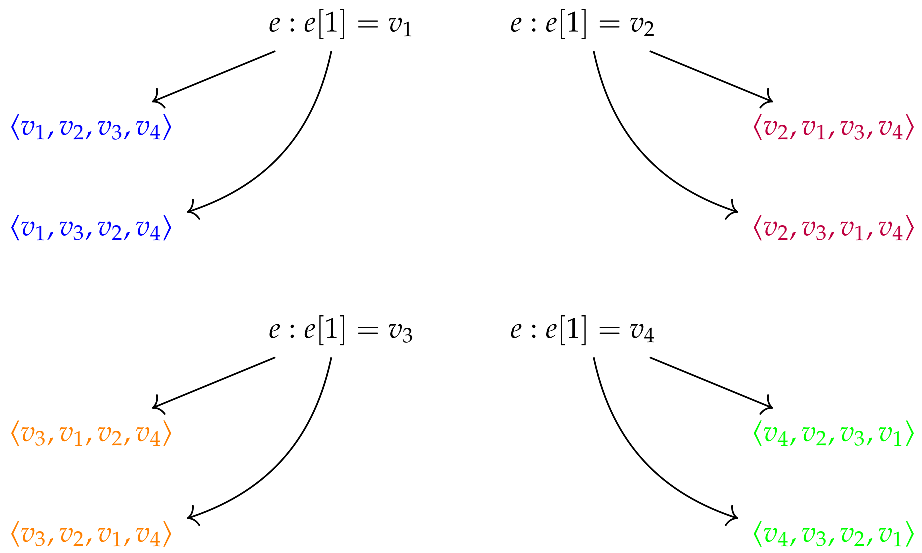


Figure 2.6: The picture of a $LBC_{\Box}$-valuation resulting from a restricted spherical model

Another connection between restricted spherical models and their induced $LBC_{\Box}$-valuations will prove essential: recall the definition of a restricted canonical spherical model $\mathcal{M}^{\mathfrak{C}} = \langle W^{\mathfrak{C}}, \mathcal{S}^{\mathfrak{C}}, \vDash^{\mathfrak{C}} \rangle$ and its induced $LBC_{\Box}$-valuation $\langle Perm(Val_{CPL}), R, \vDash \rangle$. Recall the natural selection function $f^{\mathfrak{C}}$ defined in $\mathcal{M}^{\mathfrak{C}}$ according to Lemma 0.4 such that: for all $x \in Val_{CPL}$, for all $\varphi \in For_{\mathcal{L}_{\Box\to}}$,

$$f^{\mathfrak{C}}(\varphi, x) = [\varphi]^{\mathfrak{C}} \cap min^{\varphi}_{\subseteq}(\mathcal{S}^{\mathfrak{C}}(x))$$

Furthermore, consider the natural selection function f defined over $\langle Perm(Val_{CPL}), R, \vDash \rangle$ according to Definition 2.13 such that for all $\varphi \in For_{\mathcal{L}}$,

$$f([\varphi]_{\dashv\vdash_{CPL}}, e) = \{e'[i] \mid e'[i] \leq [\varphi]_{\dashv\vdash_{CPL}}, eRe', \text{ and for all } j < i, e'[j] \nleq [\varphi]_{\dashv\vdash_{CPL}}\}$$

Lastly, consider the choice function $d : Val_{CPL} \to Perm(Val_{CPL})$ defined as:

$$d(v) = e \Leftrightarrow e[1] = v$$

Namely d associates to each element $v \in Val_{CPL}$ an element e in $Perm(Val_{CPL})$ such that the first element appearing in e is exactly v. Then

Lemma 2.21. *For any restricted canonical spherical model $\mathcal{M}^{\mathfrak{C}} = \langle W^{\mathfrak{C}}, \mathcal{S}^{\mathfrak{C}}, \vDash^{\mathfrak{C}} \rangle$ and its induced $LBC_{\Box}$-valuation $\langle Perm(Val_{CPL}), R, \vDash \rangle$, we have that for all $\varphi \in For_{\mathcal{L}}$, for all $x \in Val_{CPL}$,*

$$f^{\mathfrak{C}}(\varphi, x) = f([\varphi]_{\dashv\vdash_{CPL}}, d(x))$$

Proof. The proof in included in Appendix B.4 □

The following lemma establishes a semantic correspondence between $\mathcal{L}^{\restriction}_{\Box\!\!\rightarrow}$ and $\mathcal{L}_{LBC_{\Box}}$:

Lemma 2.22. *Consider a restricted canonical spherical model $\mathcal{M}^{\mathfrak{C}} = \langle Val_{CPL}, \mathcal{S}^{\mathfrak{C}}, \vDash^{\mathfrak{C}} \rangle$ and its induced $\mathcal{L}_{LBC_{\Box}}$-valuation $\mathcal{V}^{\mathfrak{C}} = \langle Perm(Val_{CPL}), R, \vDash \rangle$. Consider the choice function $d : Val_{CPL} \to Perm(Val_{CPL})$ as defined in Lemma 2.21.*

The following holds: for all $\varphi \in For_{\mathcal{L}^{\restriction}_{\Box\!\!\rightarrow}}$, for all $x \in Val_{CPL}$,

$$\mathcal{M}^{\mathfrak{C}}, x \vDash^{\mathfrak{C}} \varphi \Leftrightarrow \mathcal{V}^{\mathfrak{C}}, d(x) \vDash \tau(\varphi)$$

Proof. By induction on the complexity of φ. We will show two cases for exemplifications

- $\varphi = p \in Var$. It is straightforward to see that $v \vDash^{\mathfrak{C}} p \Leftrightarrow d(v) \vDash (p \mid \top)$. Indeed, the first element of $d(v)$ is exactly the classical valuation v.

- $\varphi = \psi \Box\!\!\rightarrow \delta$. Assume $x \vDash^{\mathfrak{C}} \psi \Box\!\!\rightarrow \delta$ and $d(v) = e$. Consider the natural selection function $f^{\mathfrak{C}}$ defined in $\mathcal{M}^{\mathfrak{C}}$ according to Lemma 0.4. By assumption and by Lemma 0.4 we have $f^{\mathfrak{C}}(\psi, x) \subseteq [\delta]^{\mathfrak{C}}$.

  Now, consider the selection function defined over $\mathcal{V}^{\mathfrak{C}}$ according to Definition 2.13. By Lemma 2.21 and by Lemma 2.5, we have that $f([\psi]_{\dashv\vdash_{CPL}}, e) \subseteq \{v \in Val_{CPL} \mid v(\delta) = 1\}$ since $f([\psi]_{\dashv\vdash_{CPL}}, e) = f^{\mathfrak{C}}(\psi, x)$ and $[\delta]^{\mathfrak{C}} = \{v \in Val_{CPL} \mid v(\delta) = 1\}$. Hence, by Remark 2.1, $d(x) \vDash \Box(\delta \mid \psi)$ and $\tau(\varphi) = \Box(\delta \mid \psi)$.

  Similarly for the other direction.

□

2.4.3 From $LBC_{\Box}$-valuations to spherical models

In this section, we are going to show a *reverse* operation with respect to the one illustrated in the previous section. Specifically, starting from a $LBC_{\Box}$-valuation, we will show how to construct a spherical model.

Definition 2.23. *Consider a LBC$_\Box$-valuation $\mathcal{V} = \langle Perm(Val_{CPL}), R, \vDash\rangle$, and its dual canonical lewis algebra, $\langle \mathcal{C}(\mathbf{L}), \Box\rangle$. Consider a choice function $d : Val_{CPL} \to Perm(Val_{CPL})$ as in Lemma 2.21. The total spherical model induced by $\mathcal{V} = \langle Perm(Val_{CPL}), R, \vDash\rangle$ is the tuple $\mathcal{V}^{\Box\!\!\to} = \langle Val_{CPL}, \mathcal{S}, \vDash\rangle$ such that:*

- *$\mathcal{S} : Val_{CPL} \to \wp(\wp(Val_{CPL}))$ is defined as follows: for each $v \in Val$, consider the matrix induced by $R[d(v)]$ in $\mathcal{L}$, i.e. $\mathbf{R}[d(v)]$ and its (spherical) partition $\Pi = \langle \mathbf{C}_1, \dots, \mathbf{C}_k\rangle$. Then $\mathcal{S}(v) = \{C_1, C_1 \cup C_2, \dots, C_{k-1} \cup C_k\}$*

- *$\vDash Val_{CPL} \times Var$ is defined as: for all $v \in Val_{CPL}$ and $p \in Var$,*

$$v \vDash p \Leftrightarrow v(p) = 1$$

and extended to the formulas in $\mathcal{L}_{\Box\!\!\to}$ according to Definition 0.2.

Furthermore, we set $[\varphi]^{\Box\!\!\to} = \{v \in Val_{CPL} \mid v \vDash \varphi\}$

Lemma 2.23. *Consider a LBC$_\Box$-valuation, $\mathcal{V} = \langle Perm(Val_{CPL}), R, \vDash\rangle$; the total spherical Lewisian model, $\mathcal{V}^{\Box\!\!\to} = \langle Val_{CPL}, \mathcal{S}, \vDash\rangle$, induced by $\mathcal{V}$ is indeed a total spherical Lewis' model satisfying* **Centering**.

Proof. $\mathcal{V}^{\Box\!\!\to}$ satisfies **Nestedness**, by construction of $\mathcal{S}$. $\mathcal{V}^{\Box\!\!\to}$ satisfies **Centering** since for all $v \in Val_{CPL}$, $\mathcal{S}(v) = \{C_1, \dots, C_{k-1} \cup C_k\}$ is such that $C_1 = \{v\}$ by construction of $\mathbf{R}[d(V)]$ and the fact that $\mathcal{V}$ satisfies **Cen**. Moreover, $\mathcal{V}^{\Box\!\!\to}$ satisfies **Totality** since for all $\varphi \in For_{\mathcal{L}}$ such that $\vdash_{CPL} \neg\varphi$, for all $v \in Val_{CPL}$, $\bigcup \mathcal{S}(v) = Val_{CPL}$ by construction; thus, there is a $v \in \mathcal{S}(v)$ such that $v(\varphi) = 1$. ☐

Example of a spherical Lewis' model induced by a *LBC*$_\Box$-valuation

- Consider $\langle Perm(Val_{CPL}), R, \vDash\rangle$ and $\omega \in Perm(Val_{CPL})$.
- Assume $R[\omega] = \{\langle \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6\rangle, \langle \alpha_1, \alpha_3, \alpha_2, \alpha_4, \alpha_6, \alpha_5\rangle, \langle \alpha_1, \alpha_4, \alpha_2, \alpha_3, \alpha_5, \alpha_6\rangle\}$ and the induced matrix $\mathbf{R}[\omega]$ as in the following figure[1].
- The partition Π of $\mathbf{R}[\omega]$ would be $\Pi = \{\mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3\}$.
- Assume $\alpha_1^e = \omega$. Then, $\mathcal{S}(\alpha_1) = \{C_1, C_1 \cup C_2, C_1 \cup C_3\}$. Recall that by **Sphericality**, $C_1 = c_1$, $C_2 = c_2$, $C_3 = c_5$, where $\mathbf{c}_1$ is $\mathbf{C}_1$, $\mathbf{c}_2$ is the first column of $\mathbf{C}_2$, and $\mathbf{c}_5$ is the first column of $\mathbf{C}_3$. Then $\mathcal{S}(\alpha_1) = \{c_1, c_1 \cup c_2, c_2 \cup c_3\}$, as depicted in the following figure:

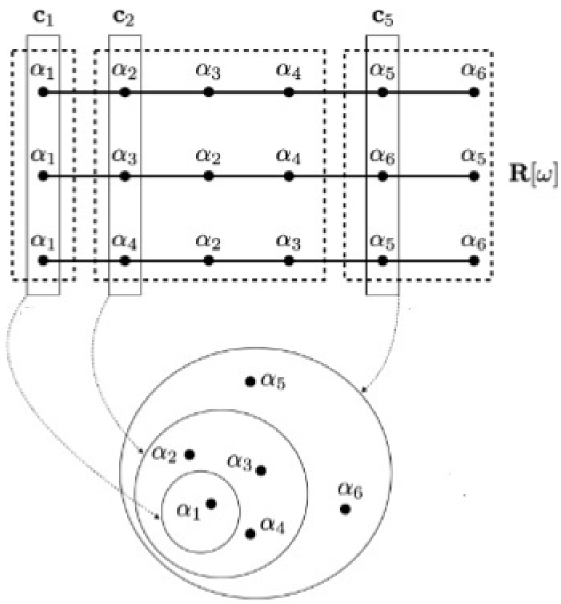


Figure 2.7: The picture illustrates how to define a sphere model out of a partition of the matrix $\mathbf{R}[\omega]$

Lemma 2.24. *Consider a* $LBC_{\Box}$*-valuation* $\mathcal{V} = \langle Perm(Val_{CPL}), R, \models\rangle$ *and its induced total spherical Lewis models* $\mathcal{V}^{\Box\!\!\rightarrow} = \langle Val_{CPL}, \mathcal{S}, \models\rangle$*.*

Consider the natural selection function $f^{\Box\!\!\rightarrow}$ *defined in* $\mathcal{V}^{\Box\!\!\rightarrow}$ *according to Lemma 0.4 such that: for all* $x \in Val_{CPL}$*, for all* $\varphi \in For_{\mathcal{L}_{\Box\!\!\rightarrow}}$*,*

$$f^{\Box\!\!\rightarrow}(\varphi, x) = [\varphi]^{\Box\!\!\rightarrow} \cap min^{\varphi}_{\subseteq}(\mathcal{S}(x))$$

Furthermore, consider the natural selection function f *defined over* $\langle Perm(Val_{CPL}), R, \models\rangle$ *according to Definition 2.13 such that for all* $\varphi \in For_{\mathcal{L}}$*,*

$$f([\varphi]_{\dashv\vdash_{CPL}}, e) = \{e'[i] \mid e'[i] \leq [\varphi]_{\dashv\vdash_{CPL}}, eRe', \text{ and for all } j < i, e'[j] \nleq [\varphi]_{\dashv\vdash_{CPL}}\}$$

Consider a choice function $d : Val_{CPL} \to Perm(Val_{CPL})$ *as in Lemma 2.21.*

We have that for all $\varphi \in For_{\mathcal{L}}$*, for all* $x \in Val_{CPL}$*,*

$$f^{\Box\!\!\rightarrow}(\varphi, x) = f([\varphi]_{\dashv\vdash_{CPL}}, d(x))$$

Proof. The proof is similar to that of Lemma 2.21. □

Now, we have all the ingredients to prove a reverse version of Lemma 2.22:

Lemma 2.25. *Consider a $LBC_{\Box}$-valuation $\mathcal{V} = \langle Perm(Val_{CPL}), R, \vDash\rangle$ and its induced total spherical Lewis' model, $\mathcal{V}^{\Box\!\!\rightarrow} = \langle Val_{CPL}, \mathcal{S}, \vDash\rangle$. Consider a choice function $d : Val \rightarrow Perm(Val_{CPL})$ defined as in Lemma 2.24.*

The following holds for all $\varphi \in For_{\mathcal{L}^{\restriction}_{\Box\!\rightarrow}}$ and all $v \in Val_{CPL}$:

$$\mathcal{V}^{\Box\!\!\rightarrow}, v \vDash \varphi \Leftrightarrow \mathcal{V}, d(v) \vDash \tau(\varphi)$$

Proof. Similar to the proof of Lemma 2.22, by induction on the complexity of φ and by employing Lemma 2.24. □

In the next section, we are going to wrap-up the results we have illustrated so far and draw some philosophical considerations about their conceptual impact.

2.4.4 A New Interpretation of Lewis counterfactuals

Recall the definition of $\mathbf{VC}^{+}$ in Notation 2.1. We have now all the ingredients to prove the following theorem:

Theorem 2.6. *For all $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}^{\restriction}_{\Box\!\rightarrow}}$,*

$$\Gamma \vdash_{\mathbf{VC}^{+}} \varphi \Leftrightarrow \tau[\Gamma] \models_{LBC_{\Box}} \tau(\varphi)$$

where $\tau[\Gamma] = \{\tau(\gamma) \mid \gamma \in \Gamma\}$

Proof.

- (⇒) By contraposition, assume $\tau[\Gamma] \not\models_{LBC_{\Box}} \tau(\varphi)$. Then, there is a $LBC_{\Box}$-valuation $\mathcal{V} = \langle Perm(Val_{CPL}), R, \vDash\rangle$ and a $e \in Perm(Val_{CPL})$ such that $\mathcal{V}, e \vDash \tau(\gamma)$ for all $\gamma \in \Gamma$, but $\mathcal{V}, e \nvDash \tau(\varphi)$. Now, consider the total spherical Lewisian model $\mathcal{V}^{\Box\!\!\rightarrow} = \langle Val_{CPL}, \mathcal{S}, \vDash\rangle$ satisfying **Centering** induced by $\mathcal{V}$. By Lemma 2.25, we have that $\mathcal{V}^{\Box\!\!\rightarrow}, d^{-1}(v) \vDash \gamma$ for all $\gamma \in \Gamma$, but $\mathcal{V}^{\Box\!\!\rightarrow}, d^{-1}(v) \nvDash \varphi$. This implies that $\Gamma \not\models_{\mathbf{VC}^{+}} \varphi$. Therefore, by Theorem 2.1, $\Gamma \nvdash_{\mathbf{VC}^{+}} \varphi$

- (⇐) By contraposition, assume $\Gamma \nvdash_{\mathbf{VC}^{+}} \varphi$. Then, by Theorem 2.1, $\Gamma \not\models_{\mathbf{VC}^{+}} \varphi$. Thus, there is a total spherical Lewis model $\mathcal{M} = \langle W, \mathcal{S}, \vDash\rangle$ satisfying **Centering** such that for some $w \in W$, $\mathcal{M}, w \vDash \gamma$ for all $\gamma \in \Gamma$, but $\mathcal{M}, w \nvDash \varphi$. By Corollary 2.1, there is a restricted canonical spherical model $\mathcal{M}^{\mathfrak{C}} = \langle Val_{CPL}, \mathcal{S}, \vDash\rangle$, and a $E^{-1}([w]_{\approx}) \in Val_{CPL}$, such that $\mathcal{M}^{\mathfrak{C}}, E^{-1}([w]_{\approx}) \vDash \gamma$ for all $\gamma \in \Gamma$, but $\mathcal{M}^{\mathfrak{C}}, E^{-1}([w]_{\approx}) \nvDash \varphi$. By Lemma 2.22, there is a $LBC_{\Box}$-valuation $\mathcal{V}^{\mathfrak{C}} = \langle Perm(Val_{CPL}), R, \vDash\rangle$ induced by $\mathcal{M}^{\mathfrak{C}}$, such that $\mathcal{V}^{\mathfrak{C}}, d(E^{-1}([w]_{\approx}) \models \tau(\gamma)$ for all $\gamma \in \Gamma$, but $\mathcal{V}^{\mathfrak{C}}, d(E^{-1}([w]_{\approx}) \not\models \varphi$. Therefore, $\tau[\Gamma] \not\models_{LBC_{\Box}} \tau(\varphi)$.

□

The above theorem establishes that a specific language fragment of the logic $\mathbf{VC}^{+}$ can be *translated* into a corresponding $LBC_{\Box}$-logical consequence. This translation converts counterfactual formulas of the form $\varphi \Box\!\!\rightarrow \psi$ into $LBC_{\Box}$-formulas of the form $\Box(\psi \mid \varphi)$, where $\Box$ is a normal modal operator obeying conditions $(L2)$ and $(L3)$ in Definition 2.12, and $(\cdot \mid \cdot)$ is a conditional operator that behave logically according to $\models_{LBC}$. As a result, the counterfactual connective can be defined in the language $\mathcal{L}_{LBC_{\Box}}$ as: $\varphi \Box\!\!\rightarrow \psi := \Box(\psi \mid \varphi)$. The resulting logic in the language $\mathcal{L}_{LBC_{\Box}}$ is essentially a slight extension of Lewis' logic of counterfactuals **VC**. Hence, according to Theorem 2.6, linguistic objects of the form $\Box(\psi \mid \varphi)$ behave logically as Lewis counterfactuals. Through the interpretation provided by τ, we can interpret a Lewis counterfactual $\varphi \Box\!\!\rightarrow \psi$ as $\Box(\psi \mid \varphi)$.

Theorem 2.6 answers the questions (L1), (L1b), (L1c), and (L2) posed in the introduction. Specifically, Theorem 2.6 offers a positive response to question (L1) regarding whether a Lewis counterfactual can be defined using other conditional operators. As mentioned earlier, we can indeed define a Lewis counterfactual within the language $\mathcal{L}_{LBC_{\Box}}$, as $\varphi \Box\!\!\rightarrow \psi := \Box(\psi \mid \varphi)$, providing a reductionist account of Lewis counterfactuals.

However, a natural question arises regarding the interpretation of the conditional operator $(\cdot \mid \cdot)$ and the modal operator $\Box$ involved in defining a Lewis counterfactual. Notably, $(\cdot \mid \cdot)$ behaves according to the axioms and rule of logic **LBC**. Furthermore, let us consider a translation $\pi : For_{\mathcal{L}^{\restriction}_{\Box\!\rightarrow}} \to For_{\mathcal{L}_{LBC}}$ inductively defined just like τ with the only difference that it translates a conditional formulas $\varphi \Box\!\!\rightarrow \psi$ into the corresponding non-modal conditional $(\psi \mid \varphi)$:

- if $\varphi \in Var$, $\pi(p) = (p \mid \top)$
- for $\varphi, \psi \in For_{\mathcal{L}}$ with $\nvdash_{CPL} \neg\varphi$, $\pi(\varphi \Box\!\!\rightarrow \psi) = (\psi \mid \varphi)$
- for $\varphi, \psi \in For_{\mathcal{L}^{\restriction}_{\Box\!\rightarrow}}$, $\pi(\varphi \wedge \psi) = \pi(\varphi) \wedge \pi(\psi)$
- for $\varphi, \psi \in For_{\mathcal{L}^{\restriction}_{\Box\!\rightarrow}}$, $\pi(\varphi \vee \psi) = \pi(\varphi) \vee \pi(\psi)$
- for $\varphi \in For_{\mathcal{L}^{\restriction}_{\Box\!\rightarrow}}$, $\pi(\neg\varphi) = \neg\pi(\varphi)$

Based on the results in (Flaminio, Godo, and Hosni 2020), it is straightforward to demonstrate that for all $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}^{\restriction}_{\Box\!\rightarrow}}$,

$$\Gamma \vdash_{\mathbf{VCS}^{+}} \varphi \Leftrightarrow \pi[\Gamma] \models_{LBC} \pi(\varphi) \Leftrightarrow \pi[\Gamma] \vdash_{\mathbf{LBC}} \pi(\varphi)$$

Namely, the above result establishes that slight extension of Stalnaker conditional logic can be interpreted as a corresponding *LBC*-consequence over the language $\mathcal{L}_{LBC}$, which in turn corresponds to logic **LBC**. It is important to recall that **VCS** coincides with Stalnaker's (1968) logic of conditionals, which in turn coincides with Adams' (1975) conditional logic over the common linguistic domain represented by $\mathcal{L}^{\restriction}_{\Box\!\!\rightarrow}$ (see Gibbard 1980), i.e. a language that does not contain nested conditionals. Therefore, linguistic objects in the form $(\psi \mid \varphi)$ behave, logically, just like Adams conditionals, and, consequently, we can identify $(\psi \mid \varphi)$ with the corresponding Adams conditional. This identification is not only justified by the fact that $(\psi \mid \varphi)$ and the corresponding Adams conditional share the same logical properties, but also by the observation that the probability of the object $(\psi \mid \varphi)$, under the distribution μ_P, is exactly the conditional probability of ψ given φ. And this fact aligns with Adams' probabilistic semantics where a conditional $A \to B$ is assigned the conditional probability of B given A. Therefore, under this Adams-like interpretation of the operator $(\cdot \mid \cdot)$, we have provided a positive answer to the question (L1b) regarding the connection between Lewis counterfactuals and Adams conditionals: a Lewis counterfactual $\varphi \Box\!\!\rightarrow \psi$ can be defined as a modality of the corresponding Adams conditional, $(\psi \mid \varphi)$, as $\varphi \Box\!\!\rightarrow \psi := \Box(\psi \mid \varphi)$. This observation also addresses question (L2) regarding the possibility of providing non-standard truth conditions for Lewis counterfactuals. Indeed, according to our new interpretation, a counterfactual $\varphi \Box\!\!\rightarrow \psi$ is true if the corresponding Adams conditional is *necessary*. In other words,

$\varphi \Box\!\!\rightarrow \psi$ is true $\Leftrightarrow$ the conditional dependence expressed by $(\psi \mid \varphi)$ holds *necessarily*

However, we have not yet provided any insights into how to interpret the term "*necessary*" in the new truth-conditions for a Lewis counterfactuals, or how to understand the modal operator $\Box$ involved in their definition. The characteristics conditions we imposed on the behavior of the box-operator, namely $(L2)$ and $(L3)$ in Definition 2.12, should provide some intuitive hints about the modality expressed by $\Box$. However, those conditions may still appear rather obscure. Let us consider $(L2)$ as an example. At the logic level, $(L2)$ implies the following equivalence: $\models_{LBC_\Box} \Box(\varphi \mid \top) \leftrightarrow (\varphi \mid \top)$. This equivalence can be interpreted as stating that, given an actual state e, $(\varphi \mid \top)$ holds at e if and only if it holds *necessarily* at e. For simplicity, using Lemma 2.8, we can identify $(\varphi \mid \top)$ with the non-conditional expression φ. Thus, $(L2)$ would imply that the non-conditional event expressed by φ holds at the actual state e if and only if it *necessarily* holds at that same state. However, the same condition does not apply to conditional events; in other words, it is not the case that $\models_{LBC_\Box} \Box(\varphi \mid \psi) \leftrightarrow (\varphi \mid \psi)$. Therefore, $(L2)$ should express something specific to non-conditional events, indicating that the proposition expressed by $\Box(\varphi \mid \top)$ is logically equivalent to the proposition expressed by $(\varphi \mid \top)$. Consequently, whatever modality $\Box$ represents, it should be irrelevant to the truth-conditions of non-conditional formulas. On the other hand, the axiom $(L3)$ is even more intricate, as it implies the validity of the following formula: $\models_{LBC_\Box} \Box(\varphi \mid \varphi \vee \psi) \vee \Box(\psi \mid \varphi \vee \psi) \vee (\Box(\delta \mid \varphi \vee \psi) \leftrightarrow \Box((\delta \mid \varphi) \wedge (\delta \mid \psi)))$. We will later attempt to provide a more transparent interpretation of this formula when introducing probability measures, as it becomes clearer under a probabilistic interpretation. For now, it is sufficient to observe that the condition expressed by $(L2)$ indeed offers some guidance on how to interpret the modality involved in the definition of a counterfactual, as the following observation will demonstrate. First, note that $(L2)$ implies that $\Box$ displays a *reflexive*-like and *transitive*-like behavior for conditionals of the form $(\varphi \mid \top)$, as it validates $\models_{LBC_\Box} \Box(\varphi \mid \top) \supset (\varphi \mid \top)$ and $\models_{LBC_\Box} \Box(\varphi \mid \top) \supset \Box\Box(\varphi \mid \top)$.

In the introduction, we mentioned the Gödel's embedding of intuitionistic logic **IL** into the modal logic **S4** (Gödel 1986; Ono 2019). Specifically, let $\mathcal{L}_{\mathbf{IL}}$ denote the standard language of intuitionistic logic, with $\rightarrow$ being the connective for intuitionistic implication, and $For_{\mathbf{IL}}$ denoting the set of formulas of $\mathcal{L}_{\mathbf{IL}}$. Similarly, let $\mathcal{L}_{\mathbf{S4}}$ denote the standard language of modal logic **S4**, and $For_{\mathbf{S4}}$ denote the set of formulas of $\mathcal{L}_{\mathbf{S4}}$. It is possible to define a translation G from $For_{\mathbf{IL}}$ into $For_{\mathbf{S4}}$ as follows:

- for $p \in Var$, $G(p) = \Box p$
- for $\varphi \wedge \psi$, $G(\varphi \wedge \psi) = G(\varphi) \wedge G(\psi)$
- for $\varphi \vee \psi$, $G(\varphi \vee \psi) = G(\varphi) \vee G(\psi)$
- for $\varphi \rightarrow \psi$, $G(\varphi \rightarrow \psi) = \Box(G(\varphi) \supset G(\psi))$

It is a well known result, due to Gödel, that

$$\vdash_{\mathbf{IL}} \varphi \Leftrightarrow \vdash_{\mathbf{S4}} G(\varphi)$$

Note that the translation G and the embedding of logic **IL** into **S4** bear some resemblances to our translation τ and Theorem 2.6. According to τ, a counterfactual conditional $\varphi \mathrel{\Box\!\!\rightarrow} \psi$ is translated into a *necessitated* conditional $\Box(\psi \mid \varphi)$, similar to how the intuitionistic conditional is translated, via G, into a *necessitated* classical material conditional. Gödel (1986) argued that the modal operator in **S4** could be interpreted as a *provability* modality, where $\Box\varphi$ expresses the fact that φ is provable. Thus, the connectives in **IL** can also be interpreted in terms of the notion of provability. For example, $\vdash_{\mathbf{IL}} \varphi \vee \psi$ would express the fact that either φ or ψ is provable, and $\vdash_{\mathbf{IL}} \varphi \to \psi$ would expresses the fact that it is provable that ψ is provable from φ, if φ is provable. Therefore, the vague analogy between the translation G and Gödel's embedding of **IL** into **S4** and our Theorem 2.6 may suggest a deeper fact: the modal operator in the $LBC_{\Box}$ language, involved in the definition of a Lewis counterfactual, might be interpreted as expressing a *provability* modality. Notably, condition $(L2)$ implies that $\Box$ behaves similar to a **S4** model operator for formulas of the form $(\varphi \mid \top)$: $\models_{LBC_{\Box}} \Box(\varphi \mid \top) \supset (\varphi \mid \top)$ (reflexivity-like behavior) and $\models_{LBC_{\Box}} \Box(\varphi \mid \top) \supset \Box\Box(\varphi \mid \top)$ (transitivity-like behavior). Moreover, in principle, there is nothing preventing us from imposing the Reflexivity and Transitivity axioms for all the formulas in $\mathcal{L}_{LBC_{\Box}}$, thus having a modality just like that of **S4**. Indeed, even under a full **S4** modality, Theorem 2.6 would still holds. Therefore, under this "provability interpretation", the truth conditions of a counterfactual would become:

$\varphi \Box\!\!\to \psi$ is true $\Leftrightarrow$ the corresponding Adams conditional $(\psi \mid \varphi)$ is *provable*

However, there might be a different alternative interpretation of the modality in **S4**, drawing again from the results connecting **IL** and modal logic. Specifically, Kripke (1965) provided a possible worlds semantics for **IL** based on Kripke models $\langle W, R, \models \rangle$, where R is reflexive and transitive (like a model for **S4**) and $\models$ is subject to a hereditary condition: for all $w, w' \in W$, if wRw' then, $w \models \varphi$ implies $w' \models \varphi$. Kripke (1965) proposes interpreting elements in a possible worlds model for **IL** as

> points in time (or "evidential situations") at which we may have various pieces of information. If at a particular point [w], we have enough information to prove a proposition A, we say that [$w \models A$]; if we lack such information, we say that [$w \not\models A$]. (Kripke 1965, p. 98)

First, observe that Kripke models for **IL** bear some resemblances to $LBC_{\Box}$-valuations. The modality $\Box$ behaves like an **S4** modality for formulas of the form $(\varphi \mid \top)$, as we previously noted, and there is a kind of hereditary condition for formulas of the form $(\varphi \mid \top)$ that also holds for $LBC_{\Box}$-valuations $\langle Perm(Val_{CPL}, R, \models) \rangle$. Specifically, if $e \models (\varphi \mid \top)$, then for all accessible e' from e, eRe', it follows that $e' \models (\varphi \mid \top)$. Kripke's interpretation of **IL**-models suggests an *epistemic* reading of intuitionistic logic and of the accessibility relation involved in its semantics. Similarly, we could interpret the modality in $LBC_{\Box}$ as epistemic. Following Kripke's suggestion, points in a $LBC_{\Box}$-valuation $\langle Perm(Val_{CPL}), R, \models \rangle$ can be understood as points in time or evidential situations. Moving from one point to another via R, we may gain or lose some pieces of information. Under this interpretation, condition ($L2$) would state that the information encoded by $(\varphi \mid \top)$ remains unchanged when transitioning between evidential situations, i.e. $\models_{LBC_\Box} \Box(\varphi \mid \top) \leftrightarrow (\varphi \mid \top)$. By identifying $(\varphi \mid \top)$ with the non-conditional expression φ, ($L2$) intuitively suggests that non-conditional pieces information remain constant across evidential situations in the model. This aligns with the fact that when evaluating a counterfactual "if Oswald hadn't killed Kennedy, someone else would have", $\varphi \boxright \psi$, as false, we take into account our actual evidence regarding Kennedy's assassination, namely that no one else beside Oswald was prepared to kill Kennedy. Therefore, given the interpretation of a counterfactual as $\varphi \boxright \psi := \Box(\psi \mid \varphi)$, when we evaluate $\Box(\psi \mid \varphi)$ with respect to an actual evidential situation e ($e \models \Box(\psi \mid \varphi)$), we check whether $(\psi \mid \varphi)$ holds in all accessible evidential situations e' that preserve our actual non-conditional evidence. According to this Kripke-style suggestion, the truth conditions of a counterfactual become:

$\varphi \boxright \psi$ is true $\Leftrightarrow$ the corresponding Adams conditional $(\psi \mid \varphi)$ is *certain*, or, *must* hold, given our available evidence

The *must* involved in the above truth-conditions should be understood as an epistemic must, similarly to the epistemic "must" interpretation of indicative conditionals presented, for instance, in Ciardelli (2021).

The two interpretations we have proposed are not mutually exclusive; in fact they can even be complementary. The modal operator in $LBC_\Box$ can be interpreted as a modality combining both epistemic and provability behavior. Under this suggestion, the truth conditions of a counterfactual could be:

$\varphi \boxright \psi$ is true $\Leftrightarrow$ the corresponding Adams conditional $(\psi \mid \varphi)$ *must be provable from our available evidence*

In Chapter 4, we will also see that our characterization of the probability of a Lewis counterfactual will provide a clearer interpretation of the modality over Lewis algebras, and, consequently, a more transparent understanding of our newly proposed truth-conditions for a Lewis counterfactual. For now, our aim was primarily to provide answers to question (L1), (L1b), (L1c) and (L2). We did not intend to offer a definitive response to the question of how to interpret hour newly proposed truth-conditions of a Lewis counterfactual; rather, our intention was to provide some hints to guide the search for such an interpretation. However, what is still missing from a technical side is an axiomatization of the logics associated to the logical consequence $\models_{LBC_\Box}$. Finding a logical system which is sound and complete for $\models_{LBC_\Box}$ is still an open problem. The most intuitive solution would be to expand the logical systems *LBC* studied by Flaminio, Godo, and Hosni (2020) with the suitable modal axioms from normal modal logic and the axioms characterizing the **Cen** and **Sph** conditions. However, proving completeness of the resulting system is not an easy task since the logic would be formulated in a very peculiar language. The main result that allows to prove completeness of the system **LBC** with respect to $\models_{\mathbf{LBC}}$ was the isomorphism between the Lindenbaum algebra of *LBC*, $\mathbf{L}_{LBC}$, and the BAC generated by the Lindenbaum algebra of classical logic, $\mathbf{L}$, that is: $\mathbf{L}_{LBC} \cong \mathcal{C}(\mathbf{L})$. However, a similar result is still missing for the modal case.

2.5 Conclusions

Wrapping up our findings, we have demonstrated how to define a Lewis counterfactual using a conditional operator from the language of $\mathcal{L}_{LBC}$ and a modal operator $\Box$, answering question (L1). As a corollary of the results in Flaminio, Godo, and Hosni (2020), the conditional operator from **LBC** can be interpreted as an Adams conditional, thereby providing a definition of Lewis counterfactuals as *necessitated* Adams conditionals, answering question (L1b). Lewis algebras and their dual Lewis frames allows for the coexistence and interaction of an Adams-like conditional $(\cdot \mid \cdot)$ and a Lewis counterfactual $\Box\!\!\rightarrow$ defined as $\varphi \Box\!\!\rightarrow \psi = \Box(\psi \mid \varphi)$, thereby providing a unified account of Adams'-like conditionals and Lewis counterfactuals, addressing question (L1c). Furthermore, we have demonstrated how our technical findings enable a different interpretation of the truth-conditions for a Lewis counterfactual, now $\Box(\varphi \mid \psi)$, compared to their standard interpretations, answering question (L2). Specifically, a counterfactual $\varphi \Box\!\!\rightarrow \psi$ is true if and only if a certain Adams-like conditional dependence between φ and ψ holds *necessarily*.

However, the results we have proven and the philosophical conclusions we have drawn so far have certain limitations. Indeed, Theorem 2.6 applies only to a restricted fragment of the language, namely the fragment $\mathcal{L}^{\restriction}_{\Box\!\rightarrow}$ where nested occurrences of counterfactual conditionals are not allowed, nor are counterfactuals with impossible antecedents. Consequently, the characterization of a Lewis counterfactual in terms of $\Box$ and $(\cdot \mid \cdot)$ is valid only for the counterfactuals in the fragment $\mathcal{L}^{\restriction}_{\Box\!\rightarrow}$. Moreover, strictly speaking, a Lewis counterfactual denotes a conditional in the logic **VC**, while our Theorem 2.6 holds for $\mathbf{VC}^{+}$. Hence, our findings would inaccurately apply to Lewis counterfactuals, as we have been asserting thus far, but rather they apply to Lewis conditionals in the logic $\mathbf{VC}^{+}$. However, the extension of **VC** with the axiom $^{+}$ is not substantial. $\mathbf{VC}^{+}$ can still conceptually serve as a suitable logic for Lewis counterfactuals, with its main distinction from **VC** lying in the validity of the *would-implies-might* schema: $\models_{\mathbf{VC}^{+}} (\varphi \Box\!\!\rightarrow \psi) \supset (\varphi \Diamond\!\!\rightarrow \psi)$, which is valid in $\mathbf{VC}^+$ but not in **VC**. Nevertheless, this schema does not pose any issue: Lewis (1973b) himself seems to endorse the idea that *would-implies-might* should be valid for counterfactuals. In **VC** models, *would-implies-might* holds true in all worlds w that contain a φ-world in their system of spheres, i.e. it is valid in the non-vacuous case. Consequently, all our findings about the characterization of Lewis counterfactuals and their new truth-conditions are not affected by the fact that our technical results are limited to $\mathbf{VC}^+$. On the other hand, the fact that our technical results are limited to the language fragment $\mathcal{L}^{\upharpoonright}_{\Box\!\!\rightarrow}$ seem to affect the generality of our logical and philosophical claims, as they would only hold for a specific class of Lewis counterfactuals, but not for the entire language. Nonetheless, in the next section, we will present a more general version of our results that apply to **VC** and all other variably strict conditional logics, as well as the full Lewis language, encompassing the entire class of Lewis counterfactuals.

Chapter 3

Counterfactuals as Definable Conditionals

In this section, we are going to show that Lewis' logics of variably strict conditionals can be *translated* into Stalnaker's logic of conditionals equipped with a normal modal operator. The translation will be carried out on a semantic level, illustrating how a Stalnakerian model, equipped with an accessibility relation, gives rise to an equivalent spherical Lewisian model, and vice versa. We will then proceed to axiomatize the resulting logic in the newly expanded language and draw some philosophical conclusions. Our findings will offer a fresh perspective on Lewis counterfactual conditionals, leading to a novel understanding of their nature.

3.1 Syntax and Semantics

The objective of this section is to provide a precise introduction to the formal languages that will be employed in this chapter. Additionally, it aims to recapitulate certain definitions and results already established in the existing literature on variably strict conditionals. While this section may seem tedious and repetitive at first glance, it is essential for the reader to pay close attention to the formal tools being introduced. These tools differ slightly from the standard model-theoretic constructions for variably strict conditionals, as they operate on various linguistic levels. Notably, both the Lewis counterfactual arrow and the Stalnaker conditional will coexist as distinct entities within our framework. To enhance readability, relevant information clarifying the notation will be placed within frames. This way, we aim to alleviate any potential confusion arising from the use of heavy notation throughout the section.

3.1.1 New Languages

Let us start with clarifying the different languages we will move through. Recall that *Var* is our enumerable set of propositional variables; in this chapter, we drop the assumption that *Var* must be finite, which was required for the results of the previous chapter. We use lowercase Latin letters $p, q, r, \ldots$ to indicate variables in *Var*. Furthermore, recall that $\mathcal{L}$ is a standard classical logical language in the signature $\neg, \wedge, \vee$, over *Var*, having the round brackets as auxiliary symbols; and $\mathcal{L}_{\Box\!\!\rightarrow}$ amounts to Lewis (1971; 1973) language for variably strict conditionals and consists in expanding $\mathcal{L}$ with the binary connective $\Box\!\!\rightarrow$. We are now going to introduce three different additional languages.

$\mathcal{L}_{>}$ corresponds to Stalnaker (1968) language for conditionals where the conditional connective $>$ is taken as primitive. Roughly, $\mathcal{L}_{>}$ is analogous to $\mathcal{L}_{\Box\!\!\rightarrow}$ with only difference that the conditional connective is expressed by the symbol $>$, instead of $\Box\!\!\rightarrow$. Formulas in $\mathcal{L}_{>}$ are defined as follows:

Definition 3.1 (Formulas of $\mathcal{L}_{>}$)**.**

- *if φ is a formula of $\mathcal{L}$, then φ is a formula of $\mathcal{L}_{>}$ too;*
- *if φ and ψ are formulas of $\mathcal{L}_{>}$ then so is $(\varphi > \psi)$;*
- *if φ, ψ are formulas of $\mathcal{L}_{>}$, then so are $\neg\varphi$, $\varphi \wedge \psi$, and $\varphi \vee \psi$*
- *nothing else is a formula of $\mathcal{L}_{>}$.*

$For_{\mathcal{L}_{>}}$ denotes the set of formulas of $\mathcal{L}_{>}$. Several non primitive connectives can be defined:

- *the conditional $(\varphi \gg \psi) := \neg(\varphi > \neg\psi)$;*
- *a box operator $\blacksquare\varphi := \neg\varphi > \varphi$;*
- *a diamond operator, $\blacklozenge\varphi := \neg\blacksquare\neg\varphi$;*
- *a comparative plausibility operator $\varphi \sqsubseteq \psi := ((\varphi \vee \psi) \gg (\varphi \vee \psi)) \supset ((\varphi \vee \psi) \gg \varphi)$*

The other language we will use, $\mathcal{L}_{>}^{\Box}$, consists in expanding Stalnaker language $\mathcal{L}_{>}$ with a modal operator $\Box$, and formulas in $\mathcal{L}_{>}^{\Box}$ are defined as follows:

Definition 3.2 (Formulas of $\mathcal{L}_{>}^{\Box}$)**.**

- *if φ is a formula of $\mathcal{L}_{>}$, then φ is a formula of $\mathcal{L}_{>}^{\Box}$ too;*
- *if φ is formulas of $\mathcal{L}_{>}^{\Box}$, then so is $\Box\varphi$;*

- *if φ, ψ are formulas of $\mathcal{L}^{\Box}_{>}$, then so are $\neg\varphi$, $\varphi \wedge \psi$, and $\varphi \vee \psi$*
- *nothing else is a formula of $\mathcal{L}^{\Box}_{>}$.*

$For_{\mathcal{L}^{\Box}_{>}}$ denotes the set of formulas of $\mathcal{L}^{\Box}_{>}$. The counterfactual conditional can be defined within $\mathcal{L}^{\Box}_{>}$ as follows: $\varphi \Box\!\!\rightarrow^{G} \psi := \Box(\varphi > \psi)$ (we use the superscript ^{G} to distinguish the derived counterfactual connective from Lewis primitive counterfactual connective).

It is interesting to point out that different modalities coexist in the language $\mathcal{L}^{\Box}_{>}$: the primitive modality $\Box$ and the modality defined by $>$, i.e. $\blacksquare$. Moreover, the modality defined via $\Box\!\!\rightarrow^{G}$, i.e. $\boxdot^{G}\varphi := \neg\varphi \Box\!\!\rightarrow^{G} \psi$ can be considered too. In general, all the derived connectives that can be defined via $\Box\!\!\rightarrow$ in $\mathcal{L}_{\Box\rightarrow}$ can analogously be defined via $\Box\!\!\rightarrow^{G}$ in $\mathcal{L}^{\Box}_{>}$. To avoid confusion, we can use the superscript ^{G} on these derived connectives in $\mathcal{L}^{\Box}_{>}$ to distinguish them from those of $\mathcal{L}_{\Box\rightarrow}$, for instance $\varphi \Diamond\!\!\rightarrow^{G} \psi := \neg(\varphi \Box\!\!\rightarrow^{G} \psi)$. Thus, the language $\mathcal{L}^{\Box}_{>}$ also includes the modality $\boxdot^{G}\varphi := \neg\varphi \Box\!\!\rightarrow^{G} \psi$. Consequently, different multimodal logics may be induced by this language and its natural semantics (see Smets and Velázquez-Quesada 2022 for a philosophical overview on multimodal logics).

Finally, $\mathcal{L}_{\restriction^{\Box}_{>}}$ is the "counterfactual" fragment of $\mathcal{L}^{\Box}_{>}$ where formulas are defined as follows:

Definition 3.3 (Formulas of $\mathcal{L}_{\restriction^{\Box}_{>}}$)**.**

- *if φ is a formula of $\mathcal{L}$, then φ is a formula of $\mathcal{L}_{\restriction^{\Box}_{>}}$ too;*
- *if φ and ψ are formulas of $\mathcal{L}_{\restriction^{\Box}_{>}}$ then so is $\Box(\varphi > \psi)$;*
- *if φ, ψ are formulas of $\mathcal{L}_{\restriction^{\Box}_{>}}$, then so are $\neg\varphi$, $\varphi \wedge \psi$, and $\varphi \vee \psi$*
- *nothing else is a formula of $\mathcal{L}^{\Box}_{>}$.*

The next tool already anticipates our main results, namely that Lewis counterfactuals (and variably strict conditionals in general) can be defined within the expanded Stalnakerian language $\mathcal{L}^{\Box}_{>}$:

Definition 3.4 (Translation from $\mathcal{L}_{\Box\rightarrow}$ into $\mathcal{L}^{\Box}_{>}$)**.** *We define a translation function $\sigma : For_{\mathcal{L}_{\Box\rightarrow}} \rightarrow For_{\mathcal{L}^{\Box}_{>}}$ as follows: for all $\varphi \in For_{\mathcal{L}_{\Box\rightarrow}}$,*

- *if $\varphi = p \in Var$, then $\sigma(p) = p$;*
- *if $\varphi = \neg\psi$, then $\tau(\neg\psi) = \neg\sigma(\psi)$*
- *if $\varphi = \psi \wedge \delta$, then $\sigma(\psi \wedge \delta) = \sigma(\psi) \wedge \sigma(\delta)$;*
- *if $\varphi = \psi \vee \delta$, then $\sigma(\psi \vee \delta) = \sigma(\psi) \vee \sigma(\delta)$;*

- *if* $\varphi = \psi \mathrel{\Box\!\!\rightarrow} \delta$, *then* $\sigma(\psi \mathrel{\Box\!\!\rightarrow} \delta) = \Box(\sigma(\psi) > \sigma(\delta))$

For a set $\Gamma \subseteq For_{\mathcal{L}_{\Box\rightarrow}}$*, we use* $\sigma[\Gamma]$ *in square brackets to indicate* $\sigma[\Gamma] = \{\sigma(\gamma) \mid \gamma \in \Gamma\}$

Namely, counterfactuals from $\mathcal{L}_{\Box\rightarrow}$ can be interpreted into $\mathcal{L}^{\Box}_{>}$ as *necessitated Stalnaker conditionals*. We will show some example of the translation σ:

Example of σ-translations

- $\sigma(p \vee (p \vee w)) = p \vee (p \vee q)$ (classically formulas remain invariant)
- $\sigma(p \mathrel{\Box\!\!\rightarrow} (p \vee q)) = \Box(p > (p \vee q))$
- $\sigma((p \mathrel{\Box\!\!\rightarrow} q) \mathrel{\Box\!\!\rightarrow} (q \mathrel{\Box\!\!\rightarrow} p)) = \Box(\Box(p > q) > \Box(q > p))$

Now, we are going to introduce some model-theoretic structures to interpret our newly introduced languages. First, we need to introduce the models for the language $\mathcal{L}_{>}$:

Definition 3.5. *A* spherical Stalnakerian model *is a tuple* $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$ *where*

- *W and* $\mathcal{S}$ *are defined according to Definition 0.2 with the additional* **Centering** *and* **Uniqueness** *constraints from Table 1 over the language* $\mathcal{L}_{>}$*;*
- $\vDash$ *is defined according to Definition 0.2 over the language* $\mathcal{L}_{>}$*, where* $>$ *replaces* $\Box\!\!\rightarrow$*;*
- *for all* $\varphi \in \mathcal{L}_{>}$*, for all* $w \in W$*,* $min^{\varphi}_{\subseteq}(\mathcal{S}(w))$ *is defined as in Definition 0.2.*

Roughly, the only difference with respect to Lewisian models is that **Limit Assumption**, **Centering**, **Uniqueness**, *and the semantic clauses for* $\vDash$ *are formulated with respect to formulas in* $\mathcal{L}_{>}$.

Moreover, for $\varphi \in For_{\mathcal{L}_{>}}$ *we set* $[\varphi] = \{w \in W \mid w \vDash \varphi\}$

A spherical Stalnakerian model amounts to a β-model for the logic **C2** defined by Lewis (1971). As the reader may observe, spherical Stalnakerian models share the same underlying structure with spherical Lewisian models. As we can see, the systems of spheres in the spherical Stalnakerian models instantiate the properties of a system of spheres in a spherical Lewisian model. However, they serve to interpret two different languages, respectively.

In order to simplify the truth-conditions of Stalnakerian conditionals, we will show an analogous of Lemma 0.1 for spherical Stalnakerian models:

Lemma 3.1. *For any* spherical Stalnakerian model $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$, *for all* $w \in W$, *for all* $\varphi, \psi \in For_{\mathcal{L}_>}$, *the following holds:*

$$w \vDash \varphi > \psi \Leftrightarrow ([\varphi] \cap min^{\varphi}_{\subseteq}(\mathcal{S}(w))) \subseteq [\psi]$$

Proof. Analogous to the proof of Lemma 0.1 with the only linguistic difference that $\Box\!\!\rightarrow$ is replaced by $>$. □

Logical consequence over spherical Stalnakerian models is defined as usual. However, we adopt a slightly different notation aligning with the notation in Lewis (1971), where **C2** denotes Stalnaker's conditional logic.

Notation 3.1.

- *Let* **C2** *denotes the class of spherical Stalnakerian models.*

For $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}_>}$,

$\Gamma \models_{\mathbf{C2}} \varphi \quad \Leftrightarrow \quad$ *for all spherical Stalnakerian models* $\mathcal{M}$, *for all* w *in* $\mathcal{M}$, *if* $\mathcal{M}, w \vDash \gamma$ *for all* $\gamma \in \Gamma$, *then* $\mathcal{M}, w \vDash \varphi$

It is straightforward to see that from Theorem 0.1 it follows that $\models_{\mathbf{C2}}$ logical consequence coincide with the logic **VCS** formulated over the language $\mathcal{L}_>$ by replacing $\Box\!\!\rightarrow$ with $>$:

Corollary 3.1. *Let* $\vdash_{\mathbf{VCS}^>}$ *be the logic induced by the* **VCS** *system in Definition 0.4 formulated in the language* $\mathcal{L}_>$ *by replacing* $\Box\!\!\rightarrow$ *with* $>$. *Then the following holds: for all* $\Gamma \cup \varphi \subseteq For_{\mathcal{L}_>}$,

$$\Gamma \vdash_{\mathbf{VCS}^>} \varphi \Leftrightarrow \Gamma \models_{\mathbf{C2}} \varphi$$

In the next section, we are going to introduce alternative model-theoretic structures to interpret the more complex language $\mathcal{L}^{\Box}_>$.

3.1.2 Spherical Kripke Models

Recall Notation 2.3. The fundamental structure of our semantic framework is that of a spherical Kripke model. This kind of models result from merging together a Kripke frame with a spherical Stalnakerian model over a common domain of possible worlds. Before providing the definition, we need some preparatory steps.

Definition 3.6. *A modal spherical Stalnakerian model is a tuple* $\langle W, \mathcal{S}, R, f \ \vDash \rangle$ *where*

- *$\langle W, R\rangle$ is a Kripke frame*
- *$f : For_{\mathcal{L}^{\Box}_{>}} \to \wp(W)$ is function defined as follows:*
    - *for all $\varphi \in For_{\mathcal{L}^{\Box}_{>}}$, for all $w \in W$, $f(\varphi, w) = \bigcup_{v \in R[w]} ([\varphi] \cap min^{\varphi}_{\subseteq}(\mathcal{S}(v)))$*

      Intuitively, $f(\varphi, w)$ selects, for all the accessible worlds v from w, the closest φ-worlds to v.
- *$\langle W, \mathcal{S}, \vDash\rangle$ is a Stalnakerian model with the only additional feature that $\vDash$ is extended to the whole set $For_{\mathcal{L}^{\Box}_{>}}$, namely the extended Stalnakerian language $\mathcal{L}^{\Box}_{>}$ can be interpreted according to the clauses in Definition 3.5 and the following additional clause:*

$$w \vDash \Box\varphi \quad \Leftrightarrow \quad \text{for all } v : wRv, v \vDash \varphi$$

Moreover, the following technical device will prove useful:

Notation 3.2.

Given a modal spherical Stalnakerian model $\langle W, \mathcal{S}, R, f \vDash\rangle$, $\equiv \subseteq W \times W$ is an equivalence relation defined as follows:

$$\equiv = \{(w, v) \mid w, v \in W \text{ and for all } \varphi \in For_{\mathcal{L}_{\restriction^{\Box}_{>}}}, w \vDash \varphi \Leftrightarrow v \vDash \varphi\}$$

namely $w \equiv v$ iff they force exactly the same formulas in the fragment $For_{\mathcal{L}_{\restriction^{\Box}_{>}}}$

Definition 3.7. *A* spherical Kripke model *is a modal Stalnakerian model of form $\langle W, R, \mathcal{S}, f, \vDash\rangle$ where f is subject to the following constraint (recall Notation 3.2 and Notation 2.3):*

- ***sphericality** of f: for all $\varphi, \psi, \delta \in For_{\mathcal{L}_{\restriction^{\Box}_{>}}}$, for all $w \in W$, either $f(\varphi \vee \psi, w)_{/\equiv} \subseteq [\varphi]_{\equiv}$, or $f(\varphi \vee \psi, w)_{/\equiv} \subseteq [\psi]_{/\equiv}$ or $f(\varphi \vee \psi, w)_{/\equiv} = f(\varphi, w)_{/\equiv} \cup f(\varphi, w)_{/\equiv}$.*

Moreover we define the proposition expressed by a formula φ as:

$$[\varphi] = \{w \in W \mid w \vDash \varphi\}$$

and $\mathcal{M} \vDash \varphi$ means that for all $w \in W$, $w \vDash \varphi$.

Roughly, a spherical Kripke model is a modal Stalnakerian model with an additional constraint on the selection function f. Observe that the function f is subject to a similar constraint as the selection function in functional Lewisian models (Definition 0.3). On the other hand, the extended semantic clauses allow us to interpret all the formulas of $\mathcal{L}^{\square}_{>}$. Furthermore, we can show that the f in a spherical Kripke model behaves just like the selection function in a functional Lewisian model over the counterfactual fragment of $\mathcal{L}^{\square}_{>}$. More precisely, the following results holds:

Lemma 3.2. *let* $\equiv$ *be defined as in Definition 3.7; in every spherical Kripke model* $\langle W, R, \mathcal{S}, f, \vDash \rangle$, f *satisfies the following constraints: for all* $w \in W$, *for all* $\varphi, \psi \in For_{\mathcal{L}^{\square}_{\restriction >}}$, ,

(i) $f(\varphi, w)_{/\equiv} \subseteq [\varphi]_{/\equiv}$

(ii) if $f(\varphi, w)_{/\equiv} \subseteq [\psi]_{/\equiv}$ *and* $f(\psi, w)_{/\equiv} \subseteq [\varphi]_{/\equiv}$, *then* $f(\varphi, w)_{/\equiv} = f(\psi, w)_{/\equiv}$

(iii) either $f(\varphi \vee \psi, w)_{/\equiv} \subseteq [\varphi]_{/\equiv}$, *or* $f(\varphi \vee \psi, w)_{/\equiv} \subseteq [\psi]_{/\equiv}$ *or* $f(\varphi \vee \psi, w)_{/\equiv} = f(\varphi, w)_{/\equiv} \cup f(\varphi, w)_{/\equiv}$.

Proof.

(i) By Definition 3.7, for all $\varphi \in For_{\mathcal{L}^{\square}_{>}}$, and for all $x \in W$, $f(\varphi, x) = \bigcup\limits_{v \in R[w]} ([\varphi] \cap min^{\varphi}_{\subseteq}(\mathcal{S}(w))) \subseteq [\varphi]$. Hence, clearly, this relation is preserved under quotient for the fragment $\mathcal{L}^{\square}_{\restriction >}$.

(ii) For arbitrary $\varphi \in For_{\mathcal{L}^{\square}_{\restriction >}}$ and $w \in W$, assume $f(\varphi, w)_{/\equiv} \subseteq [\psi]_{/\equiv}$ and $f(\psi, w)_{/\equiv} \subseteq [\psi]_{/\equiv}$. Then, by Definition 3.7, $(\bigcup\limits_{v \in R[v]} [\varphi] \cap min^{\varphi}_{\subseteq}(\mathcal{S}(v)))_{/\equiv} \subseteq [\psi]_{/\equiv}$ and $(\bigcup\limits_{v \in R[w]} [\psi] \cap min^{\psi}_{\subseteq}(\mathcal{S}(v)))_{/\equiv} \subseteq [\varphi]_{/\equiv}$. This means that for all $v \in W$ such that wRv, $([\varphi] \cap min^{\varphi}_{\subseteq}(\mathcal{S}(v)))_{/\equiv} \subseteq [\psi]_{/\equiv}$ and $([\psi] \cap min^{\psi}_{\subseteq}(\mathcal{S}(v)))_{/\equiv} \subseteq [\varphi]_{/\equiv}$. Now, it is easy to see that for all $v \in W$ such that wRv, $[\varphi] \cap min^{\varphi}_{\subseteq}(\mathcal{S}(v)) \subseteq [\psi]$ and $[\psi] \cap min^{\psi}_{\subseteq}(\mathcal{S}(v)) \subseteq [\varphi]$. Indeed, if $[\varphi] \cap min^{\varphi}_{\subseteq}(\mathcal{S}(v)) \nsubseteq [\psi]$ then clearly $([\varphi] \cap min^{\varphi}_{\subseteq}(\mathcal{S}(v)))_{/\equiv} \nsubseteq [\psi]_{/\equiv}$. Now, we can reason analogously to the proof of Lemma 0.4 to show that for all $v \in W$ such that wRv, $[\varphi] \cap min^{\varphi}_{\subseteq}(\mathcal{S}(v)) = [\psi] \cap min^{\psi}_{\subseteq}(\mathcal{S}(v))$. So, $f(\varphi, w) = \bigcup\limits_{v \in R[w]} [\varphi] \cap min^{\varphi}_{\subseteq}(\mathcal{S}(v)) = \bigcup\limits_{v \in R[w]} [\psi] \cap min^{\psi}_{\subseteq}(\mathcal{S}(v)) = f(\psi, w)$; and clearly, this relation is preserved modulo $\equiv$.

(iii) Clearly follows from the sphericality condition in Definition 3.7 of spherical Kripke model.

□

Example of a spherical Kripke model

- Roughly, a spherical Kripke model results from merging together a Kripke frame and a spherical Stalnakerian model over a common possible worlds domain.
- Consider a spherical Kripke model $\langle W, \mathcal{S}, R, f, \vDash \rangle$ where:
  - the systems of sphere are:
    * $\mathcal{S}(w_1) = \{\{w_1\}, \{w_1, w_2\}, \{w_1, w_2, w_3\}, \{w_1, w_2, w_3, w_4\}$
    * $\mathcal{S}(w_5) = \{\{w_5\}, \{w_5, w_2\}, \{w_5, w_2, w_3\}, \{w_5, w_2, w_3, w_4\}$
    * $\mathcal{S}(w_6) = \{\{w_6\}, \{w_6, w_2\}, \{w_6, w_2, w_3\}, \{w_6, w_2, w_3, w_4\}$
    * $\mathcal{S}(w_7) = \{\{w_7\}, \{w_7, w_2\}, \{w_7, w_2, w_3\}, \{w_7, w_2, w_3, w_4\}$
    * $\mathcal{S}(w_2) = \{\{w_2\}\}$
    * $\mathcal{S}(w_3) = \{\{w_3\}\}$
    * $\mathcal{S}(w_4) = \{\{w_4\}\}$
  - the accessibility relations are:
    * $R[w_1] = \{w_5, w_6, w_7\}$
    * $R[w_2] = \{w_2\}$, $R[w_3] = \{w_3\}$, $R[w_4] = \{w_4\}$, $R[w_5] = \{w_5\}$, $R[w_6] = \{w_6\}$, $R[w_7] = \{w_7\}$
  - and $\vDash$ is such that $w_5 \equiv w_6 \equiv w_7$ where $\equiv$ is defined as in Definition 3.7.

    This ensures that **Sphericality** for f holds.

In a picture:

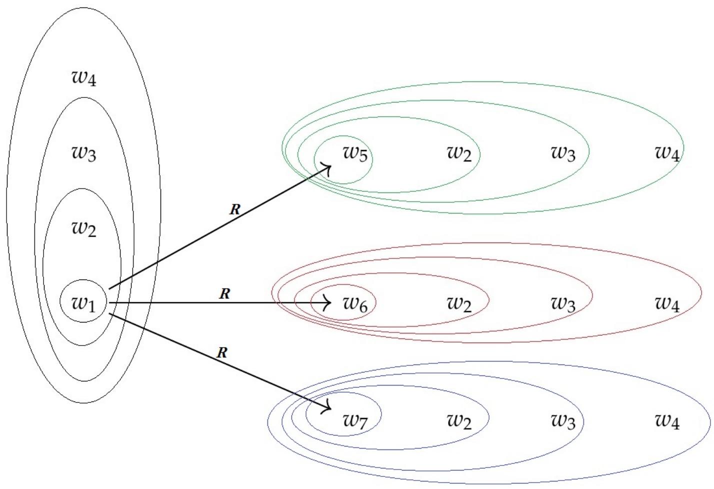


Figure 3.1: The pictures depicts and example of a spherical Kripke model

Similarly to the case of spherical Lewisian models and to Kripke frames, several classes of spherical Kripke models can be defined using different constraints on the accessibility relation R, on the systems of spheres $\mathbf{S}$, or on the defined selection function f. However, observe that the selection function f is defined via R and $\mathcal{S}$. Hence constraints on f may also implicitly encode constraints on R and $\mathcal{S}$. For the purpose of this thesis, we will consider only those constraints on the f that mirror Lewis' constraints on the selection function in Table 2. In particular, all the constraints together with their characteristic axioms are listed in the following table. It is important to note that the characteristic axiom schemata are meant to range only over formulas in the fragment $\mathcal{L}_{\restriction^{\Box}_{>}}$, and not over all formulas in $\mathcal{L}^{\Box}_{>}$.

	spherical Kripke models	
	Condition	Axiom (in $\mathcal{L}_{\restriction^{\Box}_{>}}$) for all $\varphi, \psi \in For_{\mathcal{L}_{\restriction^{\Box}_{>}}}$
(N) Normality	for some $\varphi \in For_{\mathcal{L}_{\restriction^{\Box}_{>}}}$, $f(\varphi, w)_{/\equiv} \neq \emptyset$	$\Box\varphi \supset \Diamond\varphi$
(T) Total Reflexivity	for some $\varphi \in For_{\mathcal{L}_{\restriction^{\Box}_{>}}}$, $[w]_{\equiv} \in f(\varphi, w)_{\equiv}$	$\Box(\neg\varphi > \varphi) \supset \varphi$
(W) Weak Centering	if $w \in [\varphi]$, then $[w]_{\equiv} \in f(\varphi, w)_{/\equiv}$	$\Box\varphi \supset \varphi$
(C) Centering	if $w \in [\varphi]$, then $f(\varphi, w)_{/\equiv} = \{[w]_{\equiv}\}$	$\Box\varphi \leftrightarrow \varphi$
(S) Uniqueness	$\lvert f(\varphi, w)_{/\equiv}\rvert \leq 1$	$\Diamond\varphi \supset \Box\varphi$
(U-) Local Uniformity	if $v \in \bigcup_{\varphi \in For_{\mathcal{L}_{\restriction^{\Box}_{>}}}} f(\varphi, w)$, then $\bigcup_{\varphi \in For_{\mathcal{L}_{\restriction^{\Box}_{>}}}} f(\varphi w)_{/\equiv} = \bigcup_{\varphi \in For_{\mathcal{L}_{\restriction^{\Box}_{>}}}} f(\varphi, v)_{/\equiv}$	$\Diamond\blacklozenge\varphi \supset \Box\blacksquare\Diamond\blacklozenge\varphi + \Box\blacksquare\varphi \supset \Box\blacksquare\Box\blacksquare\varphi$
(U) Uniformity	for all $w, v \in W$, $\bigcup_{\varphi \in For_{\mathcal{L}_{\restriction^{\Box}_{>}}}} f(\varphi, w)_{/\equiv} = \bigcup_{\varphi \in For_{\mathcal{L}_{\restriction^{\Box}_{>}}}} f(\varphi, v)_{/\equiv}$	$\Diamond\blacklozenge\varphi \supset \Box\blacksquare\Diamond\blacklozenge\varphi + \Box\blacksquare\varphi \supset \Box\blacksquare\Box\blacksquare\varphi$
(A-) Local Absoluteness	if $v \in \bigcup_{\varphi \in For_{\mathcal{L}_{\restriction^{\Box}_{>}}}}$, then for all $\varphi \in For_{\mathcal{L}_{\restriction^{\Box}_{>}}}$, $f(\varphi, w)_{/\equiv} = f(\varphi, v)_{/\equiv}$	$(\varphi \preccurlyeq^{G} \psi) \supset \Box\blacksquare(\varphi \preccurlyeq^{G} \psi) + \neg(\psi \preccurlyeq^{G} \varphi) \supset \Box\blacksquare\neg(\psi \preccurlyeq^{G} \varphi)$
(A) Absoluteness	for all $w, v \in W$, $f(\varphi, v)_{/\equiv} = f(\varphi, v)_{/\equiv}$	$(\varphi \preccurlyeq^{G} \psi) \supset \Box\blacksquare(\varphi \preccurlyeq^{G} \psi) + \neg(\psi \preccurlyeq^{G} \varphi) \supset \Box\blacksquare\neg(\psi \preccurlyeq^{G} \varphi)$
(UT) Universality	**Uniformity + Total Reflexivity**	**Uniformity + Total Reflexivity**
(WA) Weak Triviality	$f(\varphi, w)_{/\equiv} = [\varphi]_{\equiv}$	**Weak Centering + Absoluteness**
(CA) Triviality	if $w \in [\varphi]$, then $f(\varphi, w)_{/\equiv} = \{[w]_{\equiv}\}$, otherwise $f(\varphi, w) = \emptyset$	**Centering + Absoluteness**

Table 3.1: The table schematically summarizes the structural conditions over spherical Kripke models and the corresponding characteristic axioms. The axioms schemata for Spherical Kripke Models are meant to range over all formulas φ, ψ in $\mathcal{L}_{\restriction^{\Box}_{>}}$, i.e. formulas in the counterfactual fragment of $\mathcal{L}^{\Box}_{>}$.

Lemma 3.3. *Let $\mathcal{M} = \langle W, \mathcal{S}, R, f, \vDash \rangle$ be a spherical Kripke models. We have that*

$$\text{for all } \varphi \in For_{\mathcal{L}^{\Box}_{\restriction >}}, \mathcal{M} \vDash [\text{axiom } \mathfrak{C}] \Leftrightarrow \mathcal{M} \text{ satisfies condition } \mathfrak{C}$$

For exemplification, we will show how **Centering** in Table 3.1 characterizes the class of centered spherical Kripke models:

Theorem 3.1. *Let $\mathcal{M} = \langle W, \mathcal{S}, R, f, \vDash \rangle$ be a spherical Kripke model. We have that*

$$\text{for all } \varphi \in For_{\mathcal{L}^{\Box}_{\restriction >}}, \mathcal{M} \vDash \Box\varphi \leftrightarrow \varphi \Leftrightarrow \mathcal{M} \text{ satisfies } \textbf{Centering}$$

Proof.

(i) First, we prove that for all $\varphi \in For_{\mathcal{L}^{\Box}_{\restriction >}}, \mathcal{M} \vDash \Box\varphi \leftrightarrow \varphi$ iff for all $w \in W$, for all $\varphi \in For_{\mathcal{L}^{\Box}_{\restriction >}}$, $w \vDash \varphi \Leftrightarrow R[w] \subseteq [\varphi]$, namely that every world forces a formula in $For_{\mathcal{L}^{\Box}_{\restriction >}}$ iff all its accessible worlds force that formula too.

- ($\Rightarrow$) By contraposition, suppose that for some $\varphi \in For_{\mathcal{L}^{\Box}_{\restriction >}}$, $w \vDash \varphi$ but there is a $v \in R[w]$ such that $v \nvDash \varphi$. This means that $w \vDash \varphi$ but $w \nvDash \Box\varphi$, and so $w \nvDash \Box\varphi \leftrightarrow \varphi$. Moreover, suppose that $R[w] \subseteq [\varphi]$ but $w \nvDash \varphi$, this means that $w \vDash \Box\varphi$ but $w \nvDash \varphi$, and so $w \nvDash \Box\varphi \leftrightarrow \varphi$.
- ($\Leftarrow$) Straightforward; indeed, suppose $w \vDash \varphi$, observe that $\varphi \in For_{\mathcal{L}^{\Box}_{\restriction >}}$, hence, by assumption, $R[w] \subseteq [\varphi]$, hence $w \Vdash \Box\varphi$, that is $w \vDash \Box\varphi \supset \varphi$. Suppose $w \vDash \Box\varphi$, this means that $R[w] \subseteq [\varphi]$; so, by assumption, $w \vDash \varphi$. Therefore $w \vDash \Box\varphi \supset \varphi$. So, $w \vDash \Box\varphi \leftrightarrow \varphi$.

(ii) Now we are going to prove the main claim.

- ($\Rightarrow$) Assume that for all $\varphi \in For_{\mathcal{L}^{\Box}_{\restriction >}}, \mathcal{M} \vDash \Box\varphi \leftrightarrow \varphi$; by point ($i$) above we have that for all $\varphi \in For_{\mathcal{L}^{\Box}_{\restriction >}}$, $w \vDash \varphi$ iff $R[w] \subseteq [\varphi]$. Now, assume $w \in [\varphi]$ for $\varphi \in For_{\mathcal{L}^{\Box}_{\restriction >}}$ this means that $R[w] \subseteq [\varphi]$ and moreover, by point (i) above, that $R[w] \subseteq [w]_{\equiv}$. Moreover, observe that by definition of f, $f(\varphi, w) = R[w]$, and since $R[w] \subseteq [w]_{\equiv}$, we have that $f(\varphi, w)_{/\equiv} = \{[w]_{\equiv}\}$. So, $\mathcal{M}$ satisfies centering.
- ($\Leftarrow$) Assume $\mathcal{M}$ satisfies Centering and suppose $w \vDash \varphi$ for $\varphi \in For_{\mathcal{L}^{\Box}_{\restriction >}}$. We are going to show that $R[w] \subseteq [\varphi]$. Indeed, assume the contrary, namely there is a $v \in R[w]$ such that $w \nvDash \varphi$, hence $v \notin [w]_{\equiv}$ and $[v]_{\equiv} \neq [w]_{\equiv}$. Now, consider $[\varphi \vee \neg\varphi]$, clearly $w \in [\varphi \vee \neg\varphi]$, however, $v \in f(\varphi \vee \neg\varphi, w)$ and so, clearly $[v]_{\equiv} \in f(\varphi, w)_{/\equiv}$, hence $f(\varphi, w)_{/\equiv} \neq \{[w]_{\equiv}\}$ contradicting the assumption that $\mathcal{M}$ satisfies Centering. So, it must hold that $R[w] \subseteq [\varphi]$, and so $w \vDash \Box\varphi$, therefore $w \vDash \varphi \supset \Box\varphi$. Moreover, it must also holds that if $R[w] \subseteq [\varphi]$ for $\varphi \in For_{\mathcal{L}^{\Box}_{\restriction >}}$, then $w \vDash \varphi$. Indeed, assume that for some $\varphi \in For_{\mathcal{L}^{\Box}_{\restriction >}}$, $R[w] \subseteq [\varphi]$

but $w \nvDash \varphi$, then we can reason analogously to before and show that $f(\varphi \vee \neg\varphi, w) \neq \{[w]_\equiv\}$, contradicting the assumption that $\mathcal{M}$ satisfies Centering. Hence $w \vDash \Box\varphi \supset \varphi$. So, it must hold that $\mathcal{M}$ satisfies Centering.

□

All the other characterizations follow rather straightforwardly by the interaction between R and f in a spherical Kripke model.

Moreover, the following characterization result will prove useful:

Lemma 3.4. *In any spherical Kripke model $\langle W, R, \mathcal{S}, f, \vDash\rangle$, for all $w \in W$, for all $\varphi, \psi \in For_{\mathcal{L}^{\Box}_{>}}$, the following holds:*

$$w \vDash \Box(\varphi > \psi) \quad \Leftrightarrow \quad f(\varphi, w) \subseteq [\psi]$$

Moreover, let $\equiv$ be defined as in Definition 3.7, for all $\varphi, \psi \in For_{\mathcal{L}^{\Box}_{|>}}$,

$$w \vDash \Box(\varphi > \psi) \quad \Leftrightarrow \quad f(\varphi, w) \subseteq [\psi] \quad \Leftrightarrow \quad f(\varphi, w)_{/\equiv} \subseteq [\psi]_{/\equiv}$$

Proof. For all $\varphi, \psi \in For_{\mathcal{L}^{\Box}_{>}}$, we can reason as follows:

$$
\begin{aligned}
& w \vDash \Box(\varphi > \psi) && (3.1)\\
\Leftrightarrow\ & \text{for all } v \in W \text{ such that } wRv, v \vDash \varphi > \psi && (3.2)\\
\Leftrightarrow\ & \text{for all } v \in W \text{ such that } wRv, [\varphi] \cap min^{\varphi}_{\subseteq}(\mathcal{S}(v)) \subseteq [\psi] && (3.3)\\
\Leftrightarrow\ & \bigcup_{v \in R[w]} ([\varphi] \cap min^{\varphi}_{\subseteq}(\mathcal{S}(v))) \subseteq [\psi] && (3.4)\\
\Leftrightarrow\ & f(\varphi, w) \subseteq [\psi] && (3.5)\\
\Leftrightarrow\ & f(\varphi, w)_{/\equiv} \subseteq [\psi]_{/\equiv} && (3.6)
\end{aligned}
$$

where (3.2) and (3.3) follow from semantic conditions; (3.4) from set-theoretic operations; (3.5) from definition of f in Definition 3.7. (3.6) follows from the fact that $f(\varphi, w) \subseteq [\psi] \Leftrightarrow f(\varphi, w)_{/\equiv} \subseteq [\psi]_{/\equiv}$. This last equivalence can be easily proved as follows: ($\Rightarrow$) is straightforward; ($\Leftarrow$) is straightforward by contraposition. Indeed suppose $f(\varphi, w) \nsubseteq [\psi]$, then there is a $v \in f(\varphi, w)$ such that $v \nvDash \psi$; hence there cannot be no ψ-world that is $\equiv$-equivalent to v, hence $[v]_\equiv \in f(\varphi, w)_{/\equiv}$ but $[v]_\equiv \notin [\psi]_{/\equiv}$. □

The above result shows that the truth conditions of formulas of the form $\Box(\varphi > \psi)$ resemble those of counterfactual conditionals in Definition 0.3. And indeed formulas of the form $\Box(\varphi > \psi)$ logically behave just like counterfactuals. But, before demonstrating this last claim, we need some preparatory notions and lemmas.

In the next section we are going to examine the connections between spherical Kripke models and functional Lewisian models and their associated logical consequences.

3.2 Towards a Characterization of Lewis Variably Strict Conditionals

In this section, we will see how to pass from a spherical Kripke model into a Lewisian model for variably strict conditionals and vice versa.

3.2.1 From spherical Kripke models to Functional Models

Lemma 3.5. *Any spherical Kripke model $\mathcal{M} = \langle W, R, \mathcal{S}, f, \vDash \rangle$ encodes a functional Lewisian model $\mathcal{M}_{\equiv} = \langle W_{\equiv}, f_{\equiv}, \vDash_{\equiv} \rangle$ for $\mathcal{L}_{\Box\!\!\rightarrow}$ where $\equiv$ is defined as in Definition 3.7 and*

- *$W_{\equiv} = \{[w]_{\equiv} \mid w \in W\}$, where $[w]_{\equiv}$ is the equivalence class of w modulo $\equiv$;*
- *Let $c : W_{\equiv} \to W$ be a choice function that uniquely select for every equivalence class in $W_{\equiv}$ a representative element of the that class:*

$$c([w]_{\equiv}) \in [w]_{\equiv}$$

- *$f_{\equiv} : For_{\mathcal{L}_{\Box\!\!\rightarrow}} \to \wp(W_{/\equiv})$ is defined as follows: for all $[w]_{\equiv} \in W_{/\equiv}$ and all $\varphi \in For_{\mathcal{L}_{\Box\!\!\rightarrow}}$,*

$$f_{\equiv}(\varphi, [w]_{\equiv}) = f(\sigma(\varphi), c([w]_{\equiv}))_{/\equiv} = \{[w]_{\equiv} \mid w \in f(\sigma(\varphi), c([w]_{\equiv}))\}$$

- *$\vDash_{\equiv} \subseteq W_{\equiv} \times Var$ is defined as follows: for all $p \in Var$,*

$$[w]_{\equiv} \vDash_{\equiv} p \Leftrightarrow w \vDash p$$

  and it is extended to compound formulas as in Definition 0.3

Moreover, The following holds: for all $w \in W$, for all $\varphi \in For_{\mathcal{L}_{\Box\!\!\rightarrow}}$,

$$[w]_{\equiv} \vDash_{\equiv} \varphi \Leftrightarrow c([w]_{\equiv}) \vDash \sigma(\varphi)$$

Proof.

- It is straightforward to see that $\mathcal{M}_{\equiv}$ is a functional Lewisian model since all the three conditions $(i), (ii)$ and (iii) in Definition 0.3 hold. Specifically, this follows from Lemma 3.2, by definition of $f_{\equiv}$, and by the fact that the codomain of σ is $For_{\mathcal{L}_{\dagger>}^{\Box}}$.

- We proceed by induction on φ. The cases where $\varphi = p \in Var$, $\varphi = \neg\psi$, $\varphi = \psi \wedge \delta$, $\varphi = \psi \vee \delta$ are straightforward. We show the case where $\varphi = \psi \mathrel{\Box\!\!\rightarrow} \delta$ for exemplification: consider $w \in W$, and assume $w = c([w]_\equiv)$

$$
\begin{aligned}
& \mathcal{M}_\equiv, [w]_\equiv \vDash_\equiv \psi \mathrel{\Box\!\!\rightarrow} \delta && (3.7)\\
\Leftrightarrow\quad & f_\equiv(\psi, [w]_\equiv) \subseteq \{[w]_\equiv \mid [w]_\equiv \vDash_\equiv \delta\} && (3.8)\\
\Leftrightarrow\quad & f(\tau(\psi), w)_{/\equiv} \subseteq \{[w]_\equiv \mid w \vDash \sigma(\delta)\} && (3.9)\\
\Leftrightarrow\quad & \mathcal{M}, w \vDash \Box(\sigma(\psi) > \sigma(\delta)) && (3.10)
\end{aligned}
$$

Where (3.7) follows from semantic conditions; (3.8) follows from inductive hypothesis and by definition of $\mathcal{M}_\equiv$. Indeed, by induction hypothesis, we have that $\{[w]_\equiv \mid [w]_\equiv \vDash_\equiv \delta\} = \{[w]_\equiv \mid w \vDash \sigma(\delta)\}$ and by definition of $\mathcal{M}_\equiv$ we have that $f_\equiv(\psi, w) = f(\sigma(\psi), w)_{/\equiv}$; (3.9) follows from Lemma 3.4.

$\square$

Now, we will introduce the following notation:

Notation 3.3.

- **KV** *denotes the class of all spherical Kripke models.*
- $\mathfrak{C}$ *indicates a condition or a family of conditions (possibly empty) among those in Table 3.1, i.e.* $\{\mathbf{N}, \mathbf{T}, \mathbf{W}, \mathbf{C}, \mathbf{S}, \mathbf{A}, \mathbf{U}\}$.
- **KV**$\mathfrak{C}$ *denotes the class of spherical Kripke models satisfying condition(s)* $\mathfrak{C}$
- *Logical consequence over a class of spherical Kripke models is defined as follows: for* $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}^{\Box}_{>}}$,

$\Gamma \models_{\mathbf{KV}\mathfrak{C}} \varphi \quad \Leftrightarrow \quad$ *for all spherical Kripke models* $\mathcal{M}$ *satisfying conditions* $\mathfrak{C}$, *for all* w *in* $\mathcal{M}$, *if* $\mathcal{M}, w \vDash \gamma$ *for all* $\gamma \in \Gamma$, *then* $\mathcal{M}, w \vDash \varphi$

We can show that, modulo the translation τ, the relation of logical consequence over the language $\mathcal{L}_{\Box\!\rightarrow}$, defined over a specific class of functional (or spherical) Lewisian model, is sound with respect to the logical consequence over $\mathcal{L}^{\Box}_{>}$, defined over the corresponding class of spherical Kripke models. Indeed, the functional Lewisian model induced by a spherical Kripke model preserves the structural property of the original model. In particular, the following can be shown:

Lemma 3.6. *Let $\mathfrak{C}$ be a condition or a family of conditions (possibly empty) among those in Table 3.1, (i.e. $\{\mathbf{N}, \mathbf{T}, \mathbf{W}, \mathbf{C}, \mathbf{S}, \mathbf{A}, \mathbf{U}, \}$) and let $\mathcal{M}$ be a spherical Kripke model. Then, if $\mathcal{M}$ satisfies condition $\mathfrak{C}$, then also the functional Lewisian models induced by $\mathcal{M}$, i.e. $\mathcal{M}_\equiv$, satisfies conditions $\mathfrak{C}$.*

Proof. The result follows easily from Lemma 3.5. □

Now, we have all the ingredients to prove the following result:

Theorem 3.2. *Let $\mathfrak{C}$ be a condition or a family of conditions (possibly empty) among those in Table 3.1, (i.e. $\{\mathbf{N}, \mathbf{T}, \mathbf{W}, \mathbf{C}, \mathbf{S}, \mathbf{A}, \mathbf{U}, \}$), then, for all $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}_{\Box\!\!\to}}$, the following holds:*

$$\Gamma \vdash_{\mathbf{V}\mathfrak{C}} \varphi \Rightarrow \sigma[\Gamma] \models_{\mathbf{KV}\mathfrak{C}} \sigma(\varphi)$$

Proof. By contraposition, suppose $\sigma[\Gamma] \not\models_{\mathbf{KV}\mathfrak{C}} \sigma(\varphi)$, namely there is a spherical Kripke model $\mathcal{M} = \langle W, \mathcal{S}, R, f, \vDash\rangle$ and a $w \in W$ such that $w \vDash \sigma(\gamma)$ for all $\gamma \in \Gamma$ but $w \nvDash \varphi$. Now, consider the induced functional Lewis model $\mathcal{M}_\equiv = \langle W_\equiv, f_\equiv, \vDash_\equiv\rangle$ and a function $c : W_\equiv \to W$ such that $c([w]_\equiv) = w$. Then, by Lemma 3.6, we have that $\mathcal{M}_\equiv$ is a functional Lewisian model satisfying $\mathfrak{C}$; moreover, by Lemma 3.5, we have that $\mathcal{M}_\equiv, [w]_\equiv, \vDash_\equiv \gamma$ for all $\gamma \in \Gamma$, but $\mathcal{M}_\equiv, [w]_\equiv, \nvDash_\equiv \varphi$. Therefore, $\Gamma \not\models_{\mathbf{V}^f\mathfrak{C}} \varphi$, and by Theorem 0.1, $\Gamma \nvdash_{\mathbf{V}\mathfrak{C}} \varphi$. □

As a straightforward corollary of the above theorem and Theorem 0.1, we obtain that Lewis' logic of counterfactuals **VC** can be *embedded* into the logic induced by the local consequence defined over the spherical Kripke models satisfying centering.

Corollary 3.2. *For all $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}_{\Box\!\!\to}}$,*

$$\Gamma \vdash_{\mathbf{VC}} \varphi \Rightarrow \sigma[\Gamma] \models_{\mathbf{KVC}} \sigma[\varphi]$$

However, the above result is only close to providing a *full* interpretation of Lewis counterfactuals within the language and semantic framework of $\mathcal{L}_>^\Box$. It only demonstrates that Lewis counterfactuals are captured as linguistic objects of the form $\Box(\varphi > \psi)$ within the logical consequence $\models_{\mathbf{KVC}}$. To achieve a comprehensive interpretation of Lewis counterfactuals, we also need to establish the reverse direction: $\Gamma \vdash_{\mathbf{VC}} \varphi \Leftarrow \tau[\Gamma] \models_{\mathbf{KVC}} \tau[\varphi]$. By establishing both directions, we can confidently claim that Lewis' logic of counterfactuals is entirely captured by the logical consequence $\models_{\mathbf{KVC}}$ within the language $\mathcal{L}_>^\Box$. And this would enable us to interpret Lewis counterfactuals $\varphi \Box\!\!\to \psi$ as $\Box(\varphi > \psi)$, since both expressions would exhibit exactly the same logical properties. Although Corollary 3.2 shows that objects of the form $\Box(\varphi > \psi)$

share the logical properties of corresponding counterfactuals $\varphi \boxright \psi$, there is a possibility that $\Box(\varphi > \psi)$ might possess additional properties. In other words, certain inferences involving $\Box(\varphi > \psi)$ might hold for $\models_{\mathbf{KVC}}$ but not for $\vdash_{\mathbf{VC}}$. The reverse direction of Corollary 3.2, which will be the main focus of the next section, will establish that Lewis counterfactuals also possess the same logical properties as objects like $\Box(\varphi > \psi)$. Consequently, $\Box(\varphi > \psi)$ would have precisely the same logical properties as $\varphi \boxright \psi$.

However, before proceeding, one last comment is necessary. In Table 3.1, we have opted to use the same names (e.g. Normality, Centering, etc.) as Table 1 to denote *different* structural properties of our new model-theoretic structures. This choice was made for the sake of simplicity in notation and to avoid an excessive number of labels. Even though the names are the same, it is important to recognize that the two structural properties convey different meanings, which become evident when we examine the axioms.

For instance, in the context of spherical (functional) Lewisian models, **Centering** implies that the systems of spheres (selection functions) are centered in one singleton world, as expressed by the axiom of conjunction sufficiency $(\varphi \wedge \psi) \supset (\varphi \wedge \psi)$ and material necessity $(\varphi \boxright \psi) \supset (\varphi \supset \psi)$. On the other hand, **Centering** in the context of spherical Kripke models states that the accessibility relation, systems of spheres, and valuations must interact in a specific way, as captured by the axiom schema $\Box\varphi \leftrightarrow \varphi$. This schema appears unrelated to the counterfactual conditional $(\Box(\varphi > \psi))$ in the language $\mathcal{L}^{\Box}_{>}$, and this observation warrants philosophical attention.

Remarkably, when considering the axioms, we observe a translation of properties of the counterfactual conditional in Lewis language $\mathcal{L}_{\boxright}$ into properties of modal operators within our language $\mathcal{L}^{\Box}_{>}$ (compare the axiom for **Centering**). Similarly, at the structural level, properties of the Lewisian systems of spheres are transposed into properties of an accessibility relation within our Kripke-style structures. This phenomenon seems to align with a new potential understanding of the counterfactual conditional and its truth conditions: properties associated with $\varphi \boxright \psi$ and its corresponding Lewis' models are now being expressed in terms of $\Box$ and $>$ within the associated spherical Kripke models. Furthermore, it is worth noting that the structural properties of spherical Kripke models in Table 3.1 are formulated in a somewhat cumbersome manner through the function f. Nevertheless, it is essential to recognize that f itself is defined using the accessibility relation and the underlying Stalnakerian system of spheres.

In conclusion, we have successfully demonstrated that the Lewis' logics of counterfactuals, as well as variably strict conditionals in a broader sense, can be effectively embedded into the logical consequence relation induced by the corresponding class of spherical Kripke models. Our next objective is to present a reverse direction for this result.

3.2.2 From Spheres to Radius

In this section, we will illustrate how to build a Stalnakerian spherical model out of a spherical Lewisian model. The proofs and definitions may seem intricate and rather *ad hoc*, however they rely on some clear intuitions. Specifically, the construction we are going to present reflect very closely the construction we used in Chapter 2 to pass from a spherical Lewisian model to a Lewis frame. Indeed, what follows may be regarded as an application of the same strategy to the case of spherical Kripke models. First, we introduce the key model-theoretic construction.

Definition 3.8. *Consider a finite spherical Lewisian model* $\mathcal{M} = \langle W, \mathcal{S}, \models \rangle$*. Since* W *is finite, we have that for all* $w \in W$*,* $\mathcal{S}(w)$ *is finite too. Hence, we can index the elements in* $\mathcal{S}(w)$ *from the smallest to the greatest (recall that* $\mathcal{S}(w)$ *is totally ordered by set-inclusion), namely* $\mathcal{S}(w) = \{S_1^w, S_2^w, \ldots, S_n^w\}$ *such that* $S_1^w \subseteq S_2^w \subseteq \cdots \subseteq S_n^w$*.*

Now, define the set of radius of w s follows:

$$Radius_w = Perm(S_1^w) \times Perm(S_2^w \setminus S_1^w) \times \cdots \times Perm(S_n^w \setminus S_{n-1}^w)$$

The elements of $Radius_w$ *are tuples of tuples. However, for simplicity, with a slight abuse, we use the following notation:*

- *for* $r \in Radius_w$*,* $r[i]$ *is the element of* W *appearing in the i-th position in r;*
- *for* $r \in Radius_w$*,* $|r|$ *is the number of elements of* W *appearing in r.*

The radiation of $\mathcal{M}$*, call it* $\mathcal{M}^{\mathfrak{R}}$*, is a tuple* $\mathcal{M}^{\mathfrak{R}} = \langle W^{\mathfrak{R}}, R^{\mathfrak{R}}, \mathcal{S}^{\mathfrak{R}}, f^{\mathfrak{R}}, \models^{\mathfrak{R}} \rangle$ *where*

- $W^{\mathfrak{R}} = W \cup \bigcup_{w \in W} Radius_w$*. Roughly,* $W^{\mathfrak{R}}$ *would contain all the elements of* W *plus their radius;*
- $R^{\mathfrak{R}} \subseteq W^{\mathfrak{R}} \times W^{\mathfrak{R}}$ *is defined as follows:*
  - $wR^{\mathfrak{R}}r \Leftrightarrow w \in W$ *and* $r \in Radius_w$
  - $xR^{\mathfrak{R}}y \Leftrightarrow x, y \in W^{\mathfrak{R}} \setminus W$ *and* $y \in Radius_{x[1]}$

  namely each $w \in W$ *accesses to its radius, and any radius* $r \in W^{\mathfrak{R}} \setminus W$ *accesses to the same radius as its initial element* $r[1]$*.*

- $\mathcal{S}^{\mathfrak{R}} : W^{\mathfrak{R}} \to \wp(\wp(W^{\mathfrak{R}}))$ *is constructed as follows:*
  - *for all* $r \in \bigcup_{w \in W} Radius_w$, *let* $S_i^r = \{r[j] \mid 1 < j \leq j\} \cup \{r\}$, *namely* S_i^r *is the set of elements appearing in between the second and the i-th position in r, plus r itself. Then:*

$$\mathcal{S}^{\mathfrak{R}} = \{\{r\}\} \cup \{S_i^r \mid 1 < i \leq |r|\}$$

  - *for all* $w \in W$, *let* r_w *be one arbitrarily chosen* $r \in Radius_w$, *and let* $S_i^w = \{r[j] \mid 1 < j \leq j\} \cup \{w\}$, *namely* S_i^w *is the set of elements appearing in between the second and the i-th position in r, plus w itself then*

$$\mathcal{S}^{\mathfrak{R}}(w) = \{\{w\}\} \cup (\mathcal{S}^{\mathfrak{R}}(r_w) \setminus \{\{r\}\})$$

- *for all* $x \in W^{\mathfrak{R}}$, *we define a function* $g^{\mathfrak{R}} : For_{\mathcal{L}^{\Box}_{>}} \times W^{\mathfrak{R}} \to \wp(W^{\mathfrak{R}})$ *af follows:*

$$g^{\mathfrak{R}}(\varphi, x) = [\varphi]^{\mathfrak{R}} \cap min^{\varphi}_{\subseteq}(\mathcal{S}^{\mathfrak{R}}(x))$$

  $g^{\mathfrak{R}}$ *is simply the function that associates to each world w and formula* φ, *the "closest"* φ*-world to w according to* $\mathcal{S}^{\mathfrak{R}}(w)$.

- *consequently, we have that* $f^{\mathfrak{R}} : For_{\mathcal{L}^{\Box}_{\restriction >}} \times W^{\mathfrak{R}} \to \wp(\wp(W^{\mathfrak{R}}))$ *is such that:*

$$f^{\mathfrak{R}}(\varphi, x) = \bigcup_{y \in R^{\mathfrak{R}}[x]} g^{\mathfrak{R}}(\varphi, x)$$

- $\models^{\mathfrak{R}}: W^{\mathfrak{R}} \times Var$ *is such that: for all* $p \in Var$
  - *for all* $w \in W$, $w \models p \Leftrightarrow w \models^{\mathfrak{R}} p]$
  - *for all* $r \in \bigcup_{w \in W} Radius_w$, $r \models^{\mathfrak{R}} p \Leftrightarrow r[1] \models p$

  and $\models^{\mathfrak{R}}$ *is extended to formulas in* $\mathcal{L}^{\Box}_{>}$ *according to Definition 3.7. Moreover, we set* $[\varphi]^{\mathfrak{R}} = \{x \in W^{\mathfrak{R}} \mid x \models^{\mathfrak{R}} \varphi\}$

- $Sel^{\mathfrak{R}} : For_{\mathcal{L}^{\Box}_{\restriction >}} \times \bigcup_{w \in W} Radius_w \to \wp(W^{\mathfrak{R}})$ *is a function defined as follows:*

$$Sel^{\mathfrak{R}}(\varphi, r) = \{r[i] \mid r[i] \in [\varphi]^{\mathfrak{R}} \text{ and } \forall k < i, r[k] \notin [\varphi]^{\mathfrak{R}}\}$$

  namely, $Sel(\varphi, r)$ *outputs the (singleton of) the first* φ*-world appearing r, if any.*

Roughly, in a radiation model $\mathcal{M}^{\mathfrak{R}} = \langle W^{\mathfrak{R}}, R^{\mathfrak{R}}, \mathcal{S}^{\mathfrak{R}}, f^{\mathfrak{R}}, \models^{\mathfrak{R}} \rangle$, the system of spheres of associated to a radius r is induced by the structure of r itself.

Indeed, the smallest sphere is $\{r\}$ itself; the second-smallest sphere is the one containing r and the second element appearing in r; the third-smallest sphere is the one containing r, the second and the third element appearing in r, and so on. So that $\mathcal{S}^{\mathfrak{R}}(r)$ clearly would be totally ordered by set inclusion. Unfortunately, in order to express such intuitively property we should rely on an intricate notation and construction. On the other hand, the system of spheres associated to a $w \in W$ coincides with the system of spheres associated to some $r \in Radius_w$ except for the least element, indeed $\mathcal{S}^{\mathfrak{R}}(w)$ would be centered in $\{w\}$, while $\mathcal{S}^{\mathfrak{R}}(r)$ would be centered in $\{r\}$. The construction of a radiation model will be clearer after looking at a graphical example. Intuitively, what is happening with the radiation is an extraction of the radius of each Lewisian system of spheres so to obtain a Stalnakerian model. Now, we will prove some key properties of the radiation structure.

Lemma 3.7. *Consider the radiation $\mathcal{M}^{\mathfrak{R}} = \langle W^{\mathfrak{R}}, R^{\mathfrak{R}}, \mathcal{S}^{\mathfrak{R}}, f^{\mathfrak{R}}, \vDash^{\mathfrak{R}} \rangle$ of a finite spherical Lewisian model $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$. We have that for all $r \in W^{\mathfrak{R}} \setminus W$, for all $\varphi \in For_{\mathcal{L}_{\Box}^{>}}$,*

$$\mathcal{M}^{\mathfrak{R}}, r \vDash \varphi \Leftrightarrow \mathcal{M}^{\mathfrak{R}}, r[1] \vDash \varphi$$

Proof. By induction on φ. We show just two cases for exemplification.

Base case. By definition of $\vDash^{\mathfrak{R}}$, we have that $r \in [p]^{\mathfrak{R}}$ iff $r[1] \in [p]^{\mathfrak{R}}$.

$\varphi = \Box(\psi > \delta)$. By definition of $R^{\mathfrak{R}}$, we have that $R^{\mathfrak{R}}[r] = R^{\mathfrak{R}}[r[1]]$. Hence, it clearly holds that $x \vDash \Box(\psi > \delta)$ iff $r[1] \vDash \Box(\psi > \delta)$.

□

The above result establishes in a very straightforward way that each radius in $W^{\mathfrak{R}}$ proves exactly the same counterfactual formulas as its initial elements. The next example will help clarifying the structure of a radiation model.

Example of Radiation

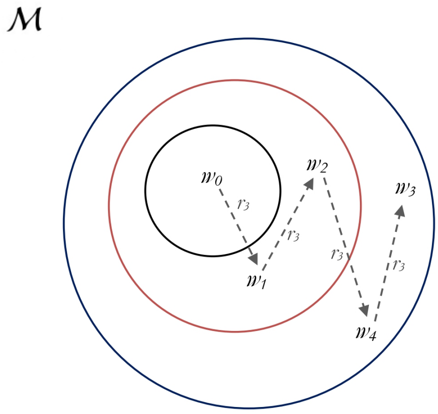
Consider a (finite) spherical Lewisian model $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$ satisfying **Centering** with:

- $w_0, w_1, w_2, w_3, w_4 \in W$
- $\mathcal{S}(w_0)$ is defined as in Figure 3.2: $\mathcal{S}(w_0) = \{\{w_0\}, \{w_0, w_1, w_2\}, \{w_0, w_1, w_2, w_3\}\}$

- We can easily see how to extract the "radius" of w_0, call them r_1, r_2, r_3, r_4, that is the elements of $Radius_w$ (see Figure 3.2). Moreover, we impose:
  - for $1 \leq i \leq 4$, $r_i \vDash^{\mathfrak{R}} p \Leftrightarrow w_0 \vDash p$
  - for $1 \leq i \leq 4$, $R^{\mathfrak{R}}[r_i] = R^{\mathfrak{R}}[w_0] = \{r_1, r_2, r_3, r_4\}$
- observe that r_1, r_2, r_3, and r_4 begin with the same possible world w_0 since $\mathcal{S}(w_0)$ is centered, and hence, r_1, r_2, r_3 and r_4 force exactly the same counterfactual formulas in $For_{\mathcal{L}_{\Box\!\!\rightarrow}}$ as w_0 according to Lemma 3.7.
- Let $\equiv$ be defined as in Defintion 3.7, namely $x \equiv y$ if and only if x and y forces the same formulas in the fragment $\mathcal{L}_{\Box\!\!\rightarrow}$, then:

$$[w_0]_{\equiv} = [r_1]_{\equiv} = [r_2]_{\equiv} = [r_3]_{\equiv} = [r_4]_{\equiv}$$

- It is also easy to see that the radiation preserves the property of **Centering**.



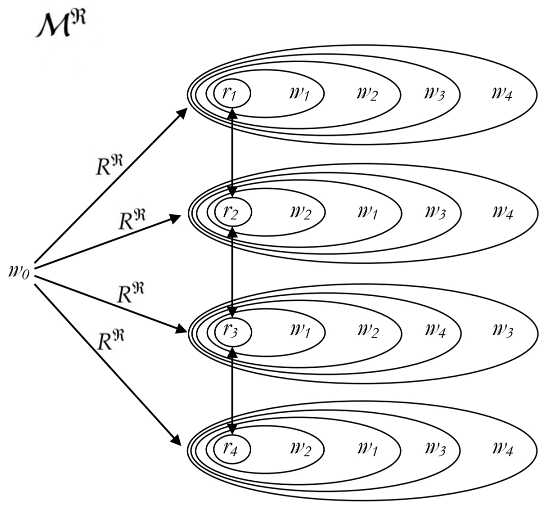
$\mathcal{S}(w_0) = \{\{w_0\}, \{w_0, w_1, w_2\}, \{w_0, w_1, w_2, w_3, w_4\}\}$

$r_1 = ((w_0), (w_1, w_2), (w_3, w_4))$

$r_2 = ((w_0), (w_2, w_1), (w_3, w_4))$

$r_3 = ((w_0), (w_1, w_2), (w_4, w_3))$

$r_4 = ((w_0), (w_2, w_1), (w_4, w_3))$



$R^{\mathfrak{R}}[w_0] = \{r_1, r_2, r_3, r_4\}$

$R^{\mathfrak{R}}[r_1] = R^{\mathfrak{R}}[r_2] = R^{\mathfrak{R}}[r_3] =$
$= R^{\mathfrak{R}}[r_4] = \{r_1, r_2, r_3, r_4\}$

$\mathcal{S}^{\mathfrak{R}}(r_1) = \{\{r_1\}, \{r_1, w_1\}, \{r_1, w_1, w_2\}, \{r_1, w_1, w_2, w_3\}\{r_1, w_1, w_2, w_3, w_4\}\}$

$\mathcal{S}^{\mathfrak{R}}(r_2) = \{\{r_1\}, \{r_1, w_2\}, \{r_1, w_2, w_1\}, \{r_1, w_2, w_1, w_3\}\{r_1, w_2, w_1, w_3, w_4\}\}$

$\mathcal{S}^{\mathfrak{R}}(r_3) = \{\{r_1\}, \{r_1, w_1\}, \{r_1, w_1, w_2\}, \{r_1, w_1, w_2, w_4\}\{r_1, w_1, w_2, w_4, w_3\}\}$

$\mathcal{S}^{\mathfrak{R}}(r_4) = \{\{r_1\}, \{r_1, w_2\}, \{r_1, w_2, w_1\}, \{r_1, w_2, w_1, w_4\}\{r_1, w_2, w_1, w_4, w_3\}\}$

Figure 3.2: A picture of the radiation construction

Now, we need to examine the structure of a radiation model in order to prove that it is indeed a spherical Kripke model. First, the following result will be useful:

Lemma 3.8. *Consider a spherical Lewisian model* $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$ *an its radiation model* $\mathcal{M}^{\Re} = \langle W^{\Re}, R^{\Re}, \mathcal{S}^{\Re}, f^{\Re}, \vDash^{\Re} \rangle$. *For each* $\varphi \in For_{\mathcal{L}_{\upharpoonright \Box >}}$, *for each* $r \in W^{\Re} \setminus W$, *let* $\equiv$ *the relation defined as in Definition 3.7: we have that*

$$g^{\Re}(\varphi, r)_{/\equiv} = Sel^{\Re}(\varphi, r)_{/\equiv}$$

recall from Definition 3.8 that $g(\varphi, r)$ *selects the closet* φ*-worlds to* r *with respect to* $\mathcal{S}^{\Re}$, *while* $Sel(\varphi, r)$ *selects the first* φ*-worlds appearing in* r.

Proof. The proof is included in Appendix C.1. The idea is that by construction of $Radius_w$ in Definition 3.8, and by the structure of r, $g(\varphi, r)$ would always coincide with $Sel(\varphi, r)$ except possibly when $Sel(\varphi, r) = r[1]$. □

Roughly, the above lemma establishes that Sel and g are essentially the same function modulo $\equiv$, namely they are essentially the same function with respect to counterfactual formulas in the fragment $\mathcal{L}_{\upharpoonright \Box >}$. Some other useful properties of the radiation directly follow from its construction. First, we need some notation recalling the notation in Chapter 2:

Notation 3.4.

- *Let* $\mathbf{Perm}(X)$ *be the matrix induced by* $Perm(X)$, *whose rows are exactly the elements of* $Perm(X)$.
- *An upper-case Latin letter, e.g.* $\mathbf{M}$, *denotes a matrix; the same letter,* M, *in italic denotes the set of the elements appearing in* $\mathbf{M}$.
- *Given a matrix* $\mathbf{M}$,
  - $\mathbf{c}_i$, *in boldface, denotes the i-th column appearing in* $\mathbf{M}$; *and* c_i *the set of elements appearing in* $\mathbf{c}_i$
  - $\mathbf{r}_i$ *denotes the i-th row (starting from above) appearing in* $\mathbf{M}$, *nd* r_i *the set of elements appearing in* $\mathbf{r}_i$
- *the operation* $\cdot$ *is defined as in Notation 2.11, i.e. for* $\mathbf{M}$ *and* $\mathbf{N}$ *matrices having the same number of rows,* $\mathbf{M} \cdot \mathbf{N}$ *is the resulting matrix obtained by juxtaposing* $\mathbf{N}$ *after* $\mathbf{M}$.

- *for a finite matrix* $\mathbf{M}$*, a partition* Π *of* $\mathbf{M}$*, is a set* $\Pi = \{\mathbf{X}_1, \ldots, \mathbf{X}_n\}$ *such that* $\mathbf{X}_1 \cdot \mathbf{X}_2 \cdot ... \cdot \mathbf{X}_n$ *results in* $\mathbf{M}$*.*

  Roughly, Π *is a division of* $\mathbf{M}$ *into sub-matrices of* $\mathbf{M}$ *having the same number of rows of* $\mathbf{M}$ *such that these sub-matrices are pairwise disjoint, i.e. they don't have any column in common.*

Lemma 3.9. *Consider the radiation* $\mathcal{M}^{\mathfrak{R}} = \langle W^{\mathfrak{R}}, R^{\mathfrak{R}}, \mathcal{S}^{\mathfrak{R}}, f^{\mathfrak{R}}, \vDash^{\mathfrak{R}} \rangle$*. For each* $x \in W^{\mathfrak{R}}$ *we have that:*

1. *for all* $w \in W^{\mathfrak{R}}$*,* $R^{\mathfrak{R}}[x] = Radius_w$ *for some* $w \in W$*;*
2. *for all* $y, z \in R^{\mathfrak{R}}[x]$*,* $|x| = |y|$
3. *Let* $\boldsymbol{Radius_w}$ *be the matrix induced by* $Radius_w$ *whose rows are exactly the elements of* $Radius_w$

   Let $\mathbf{R}^{\mathfrak{R}}[x]$ *be the matrix induced by* $R^{\mathfrak{R}}[x]$ *whose rows are exactly the elements of* $R[x]$*.*

   By point 1. above, it is immediate to see that $\mathbf{R}^{\mathfrak{R}}[x]$ *and* $\boldsymbol{Radius_w}$ *are the same matrix.*

   By Definition 3.8 we have that for $r \in Radius_w$*,*

   $$r \in Perm(S_1^w) \times Perm(S_2^w \setminus S_1^w) \times \cdots \times Perm(S_n^w \setminus S_{n-1}^w)$$

   Hence, there is a partition $\Pi = \{\boldsymbol{Perm}(S_1^r), \boldsymbol{Perm}(S_2^r), \ldots, \boldsymbol{Perm}(S_n^r)\}$ *of the matrix* $\mathbf{R}^{\mathfrak{R}}[x]$*.*

   For each $\boldsymbol{Perm}(S_i^r) \in \Pi$*, the elements in the first column of* $\boldsymbol{Perm}(S_i^r)$ *are exactly the same elements as those appearing in each row of* $\boldsymbol{Perm}(S_i^r)$*. Hence, the elements appearing in* $\boldsymbol{Perm}(S_i^r)$ *are exactly the elements appearing in its first column.*

Proof. The above lemma directly follows from the construction of a radiation model and definition of permutations. □

Example of the Matrices Induced by a Radiation Model

Consider the same example depicted above. Here, we highlighted the matrix induced by the radius/accessible elements:

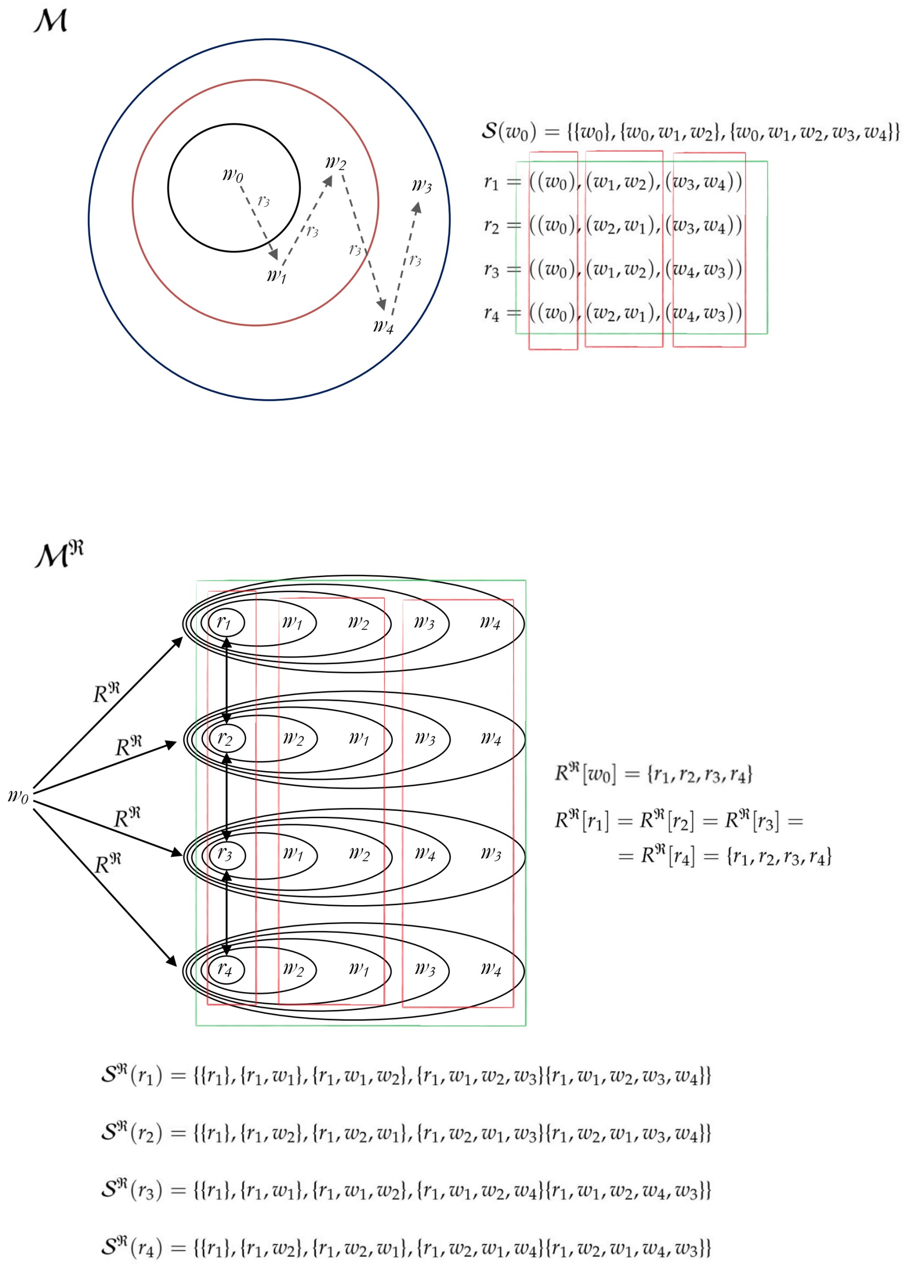


Figure 3.3: A picture of a radiation model with the induced matrix. The green and red squares represent the bigger matrix in its natural partition, respectively.

We have now all the ingredients to prove the following lemma:

Lemma 3.10. *For any fine spherical Lewisian model* $\mathcal{M} = \langle W, \mathcal{S}, \vDash\rangle$*, its radiation* $\mathcal{M}^{\mathfrak{R}} = \langle W^{\mathfrak{R}}, R^{\mathfrak{R}}, \mathcal{S}^{\mathfrak{R}}, f^{\mathfrak{R}}, \vDash^{\mathfrak{R}}\rangle$ *is a spherical Kripke model.*

Proof. We must check that $\mathcal{S}^{\mathfrak{R}}$ in $\mathcal{M}^{\mathfrak{R}}$ satisfies the constraints in Definition 3.7.

- for all $x \in W^{\mathfrak{R}}$, $\mathcal{S}^{\mathfrak{R}}(x)$ satisfies **Nestedness**. Indeed, consider $r \in W^{\mathfrak{R}} \setminus W$ and take any two S^r_i and S^r_k and assume $k < i$ without loss of generality. By construction, S^r_k contains all the elements in r from the second to the k-th position, plus r itself; on the other hand, S^r_i contains all the elements in r from the second to the i-th position, plus r itself. Hence, a fortiori, it contains all all the elements in r from the second to the k-th position, plus r itself. Hence, $S^r_i \subseteq S^r_k$. The same reasoning can be applied for $w \in W$.

- for all $x \in W^{\mathfrak{R}}$, $\mathcal{S}^{\mathfrak{R}}(x)$ satisfies **Centering** in Table 1. Immediate by construction of radiation.

- for all $x \in W^{\mathfrak{R}}$, $\mathcal{S}(x)$ satisfies **Uniqueness**. Indeed, assume for $\varphi \in For_{\mathcal{L}^{\Box}_{>}}$, $[\varphi] \cap \bigcup \mathcal{S}^{\mathfrak{R}}(x) \neq \emptyset$, then consider the minimal sphere S^x_i in $\mathcal{S}^{\mathfrak{R}}(x)$ such that $S^x_i \cap [\varphi]^{\mathfrak{R}} \neq \emptyset$, such minimal sphere clearly exists since $\mathcal{S}^{\mathfrak{R}}(x)$ is finite and totally ordered by set inclusion. Notice moreover that, by construction, $S^x_i \setminus (\bigcup_{S^x_k \subset S^x_i} S^x_k) = \{y\}$ for some $y \in W^{\mathfrak{R}}$. Moreover, we have that S^x_i is the minimal sphere such that $S^x_i \cap [\varphi]^{\mathfrak{R}} \neq \emptyset$, this means that for all the other spheres S^x_k such that $S^x_k \subset S^x_i$, $S^x_k \cap [\varphi]^{\mathfrak{R}} = \emptyset$. Hence, clearly $S^x_i \cap [\varphi] = \{y\}$.

- $f^{\mathfrak{R}}$ satisfies sphericality; i.e. for all $\varphi, \psi \in For_{\mathcal{L}^{\Box}_{|>}}$ either $f(\varphi \vee \psi, w)_{/\equiv} \subseteq [\varphi]_{/\equiv}$, or $f(\varphi \vee \psi, w)_{/\equiv} \subseteq [\psi]_{/\equiv}$ or $f(\varphi \vee \psi, w)_{/\equiv} = f(\varphi, w)_{/\equiv} \cup f(\varphi, w)_{/\equiv}$. The proof is based on a similar idea of that of Theorem 2.5. First of all, notice that $f^{\mathfrak{R}}(\varphi \vee \psi, w) \subseteq f^{\mathfrak{R}}(\varphi, w) \cup f^{\mathfrak{R}}(\psi, w)$ always holds since for all $x \in R^{\mathfrak{R}}[w]$, $min^{\varphi\vee\psi}_{\subseteq}(\mathcal{S}^{\mathfrak{R}}(x)) \cap [\varphi \vee \psi]^{\mathfrak{R}} = \{v\}$ for some $v \in W^{\mathfrak{R}}$. Where $min^{\varphi\vee\psi}_{\subseteq}(\mathcal{S}^{\mathfrak{R}}(x))$ is the minimal $\varphi \vee \psi$-permitting sphere in $\mathcal{S}^{\mathfrak{R}}(x)$, call it S^x_i. This means that for all the other $S^x_k \in \mathcal{S}^{\mathfrak{R}}(x)$ such that $S^x_k \subset S^x_i$, $S \cap [\varphi \vee \psi]^{\mathcal{R}} = \emptyset$. Hence, either $v \vDash^{\mathfrak{R}} \varphi$ or $v \vDash^{\mathfrak{R}} \psi$: if the former holds then $v \in f^{\mathfrak{R}}(\varphi, w)$ since S^x_i would also be the minimal φ-permitting sphere; and analogously if the latter holds. Hence, the claim to prove amounts to the following: either $f(\varphi \vee \psi, w)_{/\equiv} \subseteq [\varphi]_{/\equiv}$, or $f(\varphi \vee \psi, w)_{/\equiv} \subseteq [\psi]_{/\equiv}$ or $f(\varphi \vee \psi, w)_{/\equiv} \supseteq f(\varphi, w)_{/\equiv} \cup f(\varphi, w)_{/\equiv}$. In order to prove this property, we will rely on Lemma 3.9, and on Lemma 3.8. In particular, we will show the following: for all $x \in W^{\mathfrak{R}}$, for all $\varphi, \psi \in For_{\mathcal{L}_{\Box\!\!>}}$ $\bigcup_{r \in R^{\mathfrak{R}}[x]} Sel(\varphi \vee \psi, r) \subseteq [\varphi]^{\mathfrak{R}}$ or $\bigcup_{r \in R^{\mathfrak{R}}[x]} Sel(\varphi \vee \psi, r) \subseteq [\psi]^{\mathfrak{R}}$ or $\bigcup_{r \in R^{\mathfrak{R}}[x]} Sel(\varphi \vee \psi, r) \supseteq \bigcup_{r \in R[x]} Sel(\varphi, r) \cup \bigcup_{r \in R^{\mathfrak{R}}[x]} Sel(\psi, r)$. Consider the matrix $\mathbf{R}^{\mathfrak{R}}[x]$ induced by $R^{\mathfrak{R}}[x]$ and its natural partition:

$$\Pi = \{\mathbf{Perm}(S_1^w), \mathbf{Perm}(S_2^w \setminus S_1^w), \dots, \mathbf{Perm}(S_n^w)\}$$

For simplicity we will rewrite:

$$\Pi = \{\mathbf{Q}_1, \mathbf{Q}_2, \dots, \mathbf{Q}_n\}$$

where the orders of the indices mirrors the order the submatrix appear in $\mathbf{R}^{\mathfrak{R}}[x]$. For instance, if $m < l$, $\mathbf{Q}_m$ appears before $\mathbf{Q}_l$ in $\mathbf{R}^{\mathfrak{R}}[x]$. Assume $\bigcup_{r \in R^{\mathfrak{R}}[x]} Sel(\varphi \vee \psi, r) \not\subseteq [\varphi]^{\mathfrak{R}}$ and $\bigcup_{r \in R^{\mathfrak{R}}[x]} Sel(\varphi \vee \psi, r) \not\subseteq [\psi]^{\mathfrak{R}}$. This means that there are two $y, z \in R^{\mathfrak{R}}[x]$ such that $Sel(\varphi \vee \psi, y) \vDash^{\mathfrak{R}} \varphi \wedge \neg\psi$ and $Sel(\varphi \vee \psi, z) \vDash^{\mathfrak{R}} \psi \wedge \neg\varphi$, hence $y \neq z$. Moreover, by definition of *Sel*, it is the case that for some i, $y[i] = Sel(\varphi \vee \psi, y)$, and for some j, $z[j] = Sel(\varphi \vee \psi, x)$, and moreover $x[j] \neq y[i]$ since they prove different formulas by assumption. Now, we have two cases to consider:

1. $y[i], z[j]$ are in the same submatrix $\mathbf{Q}_l \in \Pi$. This means that $z[j]$ and $y[i]$ appears in each row of $\mathbf{Q}_l$ by Lemma3.9. Now, $z[j], y[i] \in Sel(\varphi \vee \psi, x)$, implies $Sel(\varphi \vee \psi, x) \subseteq Q_l$. Indeed, by way of contradiction, assume there is $u \in Sel(\varphi \vee \psi, x)$ such that $u \notin Q_m$. If u in $\mathbf{Q}_l$ for some $m > l$, by definition of *Sel*, and by Lemma 3.9, $y[i]$ or $z[j]$ would appear before u on the same row, contradicting the assumption that $u \in f(\varphi \vee \psi, x)$. Thus, assume that u in $\mathbf{Q}_o$ for some $o < l$. In this case, by Lemma 3.9, again, u appears in each row of $\mathbf{Q}_o$, and so, it must appears before $y[i]$ ($z[j]$), in the same rows as $y[i]$ ($z[j]$). This implies that $y[i] \neq u$ and $z[j] \neq u$ and $y[i], z[j] \notin Sel(\varphi \vee \psi, x)$, contradicting our assumption. So, $Sel(\varphi \vee \psi, x) \subseteq Q_l$. By an analogous reasoning and using the fact that $y[i] \vDash^{\mathfrak{R}} \varphi$ and $y[i] \nvDash^{\mathfrak{R}} \psi$, and $z[j] \vDash^{\mathfrak{R}} \psi$ but $z[j] \nvDash^{\mathfrak{R}} \varphi$, we can show that $Sel(\varphi, z) \subseteq Q_l$ and $Sel(\psi, x) \subseteq Q_l$.

   Let $\mathbf{q}_l$ be the first column of $\mathbf{Q}_l$ so that $Sel(\varphi, x) \subseteq q_l = Q_l$, $Sel(\psi, x) \subseteq q_l = Q_l$, and $Sel(\varphi \vee \psi, x) \subseteq q_l = Q_l$. Hence, $Sel(\varphi, x), Sel(\psi, x), Sel(\varphi \vee \psi, x)$ are all subsets of q_l. Therefore, it cannot exists $u \in Sel(\varphi \vee \psi, x)$ but $u \notin Sel(\varphi, x)$ and $u \notin Sel(\psi, x)$ because if $u \vDash^{\mathfrak{R}} \varphi \vee \psi$, then $u \vDash \varphi$ or $u \vDash \psi$ and $u \in c_l$. And this implies $Sel(\varphi, x) \cup Sel(\psi, x) \subseteq Sel(\varphi \vee \psi, x)$.

2. $y[i]$ and $z[j]$ are in two different sub-matrices of Π. i.e. for some $\mathbf{Y}, \mathbf{Z} \in \Pi$, $z[j]$ in $\mathbf{Z}$, and $y[i]$ in $\mathbf{Y}$. By contradiction, we show that this case cannot hold. Indeed, and without loss of generality, assume that $\mathbf{Y}$ precedes $\mathbf{Z}$ in the bigger matrix $\mathbf{R}^{\mathfrak{R}}[x]$. Then, by Lemma 3.9, $y[i]$ appears in each row of $\mathbf{Y}$ and $z[j] \notin Y$. By reasoning analogously to the case 1 above, we can show that $Sel(\varphi \vee \psi, \omega) \subseteq Y$. This leads to a contradiction with the fact that $z[j] \in Sel(\varphi \vee \psi, x)$ but $z[j] \notin Y$

Then, we have shown that for all $x \in W^{\mathfrak{R}}$, for all $\varphi, \psi \in For_{\mathcal{L}_{\restriction\overset{\Box}{>}}}$, $\bigcup_{r \in R^{\mathfrak{R}}[x]} Sel(\varphi \vee \psi, r) \subseteq [\varphi]^{\mathfrak{R}}$ or $\bigcup_{r \in R^{\mathfrak{R}}[x]} Sel(\varphi \vee \psi, r) \subseteq [\psi]^{\mathfrak{R}}$ or $\bigcup_{r \in R^{\mathfrak{R}}[x]} Sel(\varphi \vee \psi, r) = \bigcup_{r \in R^{\mathfrak{R}}[x]} Sel(\varphi, r) \cup \bigcup_{r \in R^{\mathfrak{R}}[x]} Sel(\psi, r)$. Therefore, we clearly have that $\bigcup_{r \in R^{\mathfrak{R}}[x]} Sel(\varphi \vee \psi, r)_{/\equiv} \subseteq [\varphi]^{\mathfrak{R}}$ or $\bigcup_{r \in R^{\mathfrak{R}}[x]} Sel(\varphi \vee \psi, r)_{/\equiv} \subseteq [\psi]^{\mathfrak{R}}$ or $\bigcup_{r \in R^{\mathfrak{R}}[x]} Sel(\varphi \vee \psi, r)_{/\equiv} = \bigcup_{r \in R^{\mathfrak{R}}[x]} Sel(\varphi, r)_{/\equiv} \cup \bigcup_{r \in R^{\mathfrak{R}}[x]} Sel(\psi, r)_{/\equiv}$. So, by Lemma 3.8, we have that $\bigcup_{r \in R^{\mathfrak{R}}[x]} g(\varphi \vee \psi, r)_{/\equiv} \subseteq [\varphi]^{\mathfrak{R}}$ or $\bigcup_{r \in R^{\mathfrak{R}}[x]} g(\varphi \vee \psi, r)_{/\equiv} \subseteq [\psi]^{\mathfrak{R}}$ or $\bigcup_{r \in R^{\mathfrak{R}}[x]} g(\varphi \vee \psi, r)_{/\equiv} = \bigcup_{r \in R^{\mathfrak{R}}[x]} g(\varphi, r)_{/\equiv} \cup \bigcup_{r \in R^{\mathfrak{R}}[x]} g(\psi, r)_{/\equiv}$ and so that either $f(\varphi \vee \psi, w)_{/\equiv} \subseteq [\varphi]_{/\equiv}$, or $f(\varphi \vee \psi, w)_{/\equiv} \subseteq [\psi]_{/\equiv}$ or $f(\varphi \vee \psi, w)_{/\equiv} = f(\varphi, w)_{/\equiv} \cup f(\varphi, w)_{/\equiv}$

□

Moreover, it is not difficult to show that for any property $\mathfrak{C}$ in Table 1, the radiation construction preserves $\mathfrak{C}$:

Lemma 3.11. *For any property $\mathfrak{C}$ in Table 1, given a spherical Lewis model $\mathcal{M}$ satisfying $\mathfrak{C}$, its radiation $\mathcal{M}^{\mathfrak{R}}$ satisfies the corresponding property $\mathfrak{C}$ in Table 3.1.*

Proof. We will just show one case for exemplification. Consider a finite spherical Lewisian model $\mathcal{M} = \langle W, \mathcal{S}, \rangle$ satisfying **Centering**, then its radiation $\mathcal{M}^{\mathfrak{R}} = \langle W^{\mathfrak{R}}, R^{\mathfrak{R}}, \mathcal{S}^{\mathfrak{R}}, f^{\mathfrak{R}}, \vDash^{\mathfrak{R}} \rangle$ satisfies **Centering** from Table 3.1.

Since $\mathcal{M}$ satisfies **Centering**, we have that for all $w \in W$, all the radius of $Radius_w$, according to Definition 3.8, begin with the same elements, i.e. for all $r, r' \in Radius_w$, $r[1] = r'[1]$. Consider any $x \in W^{\mathfrak{R}}$. By construction of radiation, we have that $R^{\mathfrak{R}}[x] = Radius_w$ for some $w \in W$. Now, assume $x \vDash^{\mathfrak{R}} \varphi$ for $\varphi \in For_{\mathcal{L}_{\restriction\overset{\Box}{>}}}$. Now, if $x \in W$, then, by **Centering**, $r[1] = w$ for all $r \in Radius_w$ and, by Lemma 3.7, for all $r \in Radius_w$, r forces exactly the same formulas in $For_{\mathcal{L}_{\restriction\overset{\Box}{>}}}$ as w, hence in particular φ. So, $f^{\mathfrak{R}}(\varphi, w) = Radius_w$, and so $[w]_{\equiv} = f^{\mathfrak{R}}(\varphi, x)_{/\equiv}$. Analogously if $x \in Radius_w$. □

We are now going to prove an essential lemma, namely that when passing from spherical Lewisian models to radiation models, satisfaction of counterfactual formulas is not affected:

Lemma 3.12. *For any finite Lewisian model* $\langle W, g, \vDash\rangle$*, its radiation* $\mathcal{M}^{\mathfrak{R}} = \langle W^{\mathfrak{R}}, R^{\mathfrak{R}}, S^{\mathfrak{R}}, g^{\mathfrak{R}}, f^{\mathfrak{R}}, \vDash^{\mathfrak{R}}\rangle$ *is such that for all* $\varphi \in For_{\mathcal{L}_{\Box\!\rightarrow}}$*, for all* $w \in W$*,*

$$\mathcal{M}, w \vDash \varphi \Leftrightarrow \mathcal{M}^{\mathfrak{R}}, w \vDash^{\mathfrak{R}} \sigma(\varphi)$$

Proof. By induction; the only interesting case is for $\varphi = \psi \Box\!\!\rightarrow \delta$. For simplicity, consider

$$\mathcal{S}(w) = \{S_1^w, S_2^w, \ldots, S_n^w\}$$

and the matrix induced by $\mathbf{R}^{\mathfrak{R}}[w] = \mathbf{Radius}_w$ such that its partition is

$$\Pi = \{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n\}$$

Observe that by construction for all $1 < i \leq n$, $X_1 = S_1^w$ and $X_i = S_i^w \setminus S_{i-1}^w$, thus $S_i^w = \bigcup_{1 \leq j \leq i} X_j$

($\varphi = \psi \Box\!\!\rightarrow \delta$). By contraposition, assume $\mathcal{M}, w \nvDash \psi \Box\!\!\rightarrow \delta$, this means that $[\psi] \cap \bigcup \mathcal{S}(w) \neq \emptyset$ and for all $S \in \mathcal{S}(w)$, if $S \cap [\psi] \neq \emptyset$, then $S \cap [\psi] \cap [\neg\delta] \neq \emptyset$. This implies that, for the minimal ψ-permitting sphere $S_i^w \in \mathcal{S}(w)$, there is a $v \in W$ such that $v \in S_i^w$ and $v \vDash \psi$ and $v \nvDash \delta$. Moreover, for all $u \in S_j^w$ such that $j < i$, $u \nvDash \psi$. Observe that, by construction of radiation and by Lemma 3.9, for all $X_j \in \Pi$ such that $j < i$, $X_j = S_j^w \setminus S_{j-1}^w$. Thus, by induction hypothesis, for all $u \in X_j$ such that $j < i$, $u \nvDash^{\mathfrak{R}} \sigma(\psi)$. So, by construction of the matrix, v must appear in the first column of $\mathbf{X}_i$. Hence, there must be a $r \in Radius_w$, and an index k such that $r[k] = v$ and for all $l < k$, $r[k] \nvDash \psi$. So, $Sel(\tau(\psi), r) = \{v\}$. Then, by induction hypothesis, we have that $r[k] = v \vDash^{\mathfrak{R}} \sigma(\psi)$ and $r[k] = v \nvDash^{\mathfrak{R}} \sigma(\delta)$. Thus, $Sel(\tau(\psi), r) = \{v\} \not\subseteq [\delta]^{\mathfrak{R}}$. This means that $Sel(\tau(\psi), r)_{/\equiv} \not\subseteq [\tau(\psi)]^{\mathfrak{R}}_{\equiv}$, and so, by Lemma 3.8, $g^{\mathfrak{R}}(\tau(\varphi), r)_{/\equiv} \not\subseteq [\tau(\psi)]^{\mathfrak{R}}_{\equiv}$. Since $r \in R^{\mathfrak{R}}[w]$, then this also implies that $f^{\mathfrak{R}}(\varphi, w)_{/\equiv} \not\subseteq [\tau(\psi)]^{\mathfrak{R}}_{\equiv}$, and so $w \nvDash^{\mathfrak{R}} \Box(\tau(\varphi) > \tau(\psi))$, that is $w \nvDash^{\mathfrak{R}} \tau(\varphi \Box\!\!\rightarrow \psi)$.

For the other direction, we can reason analogously to the above case but backwards. In particular, assume $w \nvDash^{\mathfrak{B}} \Box(\sigma(\psi) > \sigma(\delta))$, this means that $f^{\mathfrak{R}}(\sigma(\psi)) \not\subseteq [\sigma(\delta)]^{\mathfrak{R}}$. So, there must be a $r \in Radius_w$ such that $g^{\mathfrak{R}}(\sigma(\psi), r) \not\subseteq [\sigma(\delta)]^{\mathfrak{R}}$. Hence, it must also be the case that $g^{\mathfrak{R}}(\sigma(\psi), r)_{/\equiv} \not\subseteq [\sigma(\delta)]^{\mathfrak{R}}_{\equiv}$. So, by Lemma 3.8, $Sel(\sigma(\psi), r)_{/\equiv} \not\subseteq [\sigma(\delta)]^{\mathfrak{R}}_{\equiv}$. Thus, $Sel(\sigma(\psi), r) \not\subseteq [\sigma(\delta)]^{\mathfrak{R}}$. This means that there is a $v \in W$, such that $Sel(\sigma(\psi), r) = \{v\}$. So, $v \vDash^{\mathfrak{R}} \sigma(\psi)$ and $v \nvDash^{\mathfrak{R}} \sigma(\delta)$. So, by induction hypothesis, $v \vDash \psi$ and $v \nvDash \delta$. Moreover, $Sel(\sigma(\psi), r) = \{v\}$ implies that for some i, $r[i] = v$ and for all $k < i$, $r[k] \nvDash^{\mathfrak{R}} \sigma(\psi)$. Hence, by induction hypothesis, all $k < i$, $r[k] \nvDash^{\mathfrak{R}} \psi$. So there is a $\mathbf{X}_i \in \Pi$ such that $v \in X_i$ and for all $j < i$, $X_i \cap [\psi] = \emptyset$. Hence, S_i^w is the minimal ψ-permitting sphere in $\mathcal{S}(w)$, but $v \in S_i^w$ and $v \vDash \psi$ and $v \nvDash \delta$. Hence. $w \nvDash \psi \mathbin{\Box\!\!\rightarrow} \delta$.

□

We are now ready to prove the main theorem of this section:

Theorem 3.3. *For finite* $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}_{\Box\!\rightarrow}}$,

$$\Gamma \models_{\mathbf{V}\mathfrak{C}} \varphi \Leftarrow \sigma[\Gamma] \models_{\mathbf{KV}\mathfrak{C}} \sigma(\varphi)$$

Proof. By contraposition, assume $\Gamma \nvdash_{\mathbf{V}\mathfrak{C}} \varphi$. Then, by Theorem 0.1, $\Gamma \not\models_{\mathbf{V}\mathfrak{C}} \varphi$. By the finite model property of variably strict conditionals logics (see Lewis 1973b), there is a finite spherical Lewisian model $\mathcal{M} = \langle W, \mathcal{S}, \vDash\rangle$, satisfying $\mathfrak{C}$, and some $w \in W$, such that $w \vDash \gamma$ for all $\gamma \in \Gamma$ and $w \nvDash \varphi$. By Lemma 3.11 and Lemma 3.12, the radiation model $\mathcal{M}^{\mathfrak{R}} = \langle W^{\mathfrak{R}}, R^{\mathfrak{R}}\mathcal{S}^{\mathfrak{R}}, f^{\mathfrak{R}}, \vDash^{\mathfrak{R}}\rangle$ also satisfies $\mathfrak{C}$ and moreover, $w \vDash^{\mathfrak{R}} \sigma(\gamma)$ for all $\gamma \in \Gamma$ and $w \nvDash \sigma(\varphi)$. Hence $\sigma[\Gamma] \not\models_{\mathbf{KV}\mathfrak{C}} \sigma(\varphi)$.

□

As a corollary of the above theorem, we get:

Corollary 3.3. *For finite* $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}_{\Box\!\rightarrow}}$,

$$\Gamma \vdash_{\mathbf{VC}} \varphi \Leftarrow \sigma[\Gamma] \models_{\mathbf{KVC}} \sigma(\varphi)$$

3.2.3 A New Perspective on Lewis counterfactuals

Combining the results from Corollary 3.3 and Corollary 3.2, we arrive at the following theorem.

Theorem 3.4. *For finite* $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}^{\Box}_{>}}$,

$$\Gamma \vdash_{\mathbf{VC}} \varphi \Leftrightarrow \sigma[\Gamma] \models_{\mathbf{KVC}} \sigma(\varphi)$$

With this theorem, we have now achieved a comprehensive characterization of the entire class of Lewis counterfactuals (and variably strict conditionals) within the language $\mathcal{L}^{\Box}_{>}$. This represents a significant refinement of the findings from Chapter 2. While in the BACs framework we translated an extension of Lewis' logic of counterfactuals, i.e $\mathbf{VC}^{+}$, into the logical consequence $\models_{LBC_{\Box}}$, here we have provided a faithful translation of Lewis' logic **VC** directly into the logical consequence $\models_{\mathbf{KVC}}$. Furthermore, in contrast to the limitations of the BACs framework where we were confined to a fragment of Lewis language containing only non-nested counterfactuals with satisfiable antecedents, our Theorem 3.4 encompasses the entire Lewis language. This enhanced generality provides more comprehensive and robust answers to the questions (L1), (L1a), (L1c), and (L2) posed in the introduction.

Let us examine our results from a broader philosophical perspective. Firstly, we can express the Lewis counterfactual connective within or language $\mathcal{L}^{\Box}_{>}$ as $\varphi \mathrel{\Box\!\!\rightarrow} \psi := \Box(\varphi > \psi)$. Remarkably, the logical consequence over formulas $For_{\mathcal{L}^{\Box}_{\restriction >}}$, with respect to spherical Kripke models satisfying **Centering**, coincides precisely with Lewis' logic of counterfactual **VC**. This alignment, at the level of the logic, between objects of the form $\Box(\varphi > \psi)$ and Lewis counterfactuals $\varphi \mathrel{\Box\!\!\rightarrow} \psi$ provides support for the identification of our $\Box(\varphi > \psi)$ with the Lewis expression $\varphi \mathrel{\Box\!\!\rightarrow} \psi$. Thus, we successfully addressed question (L1) concerning the possibility of constructing a reductionist account of Lewis counterfactuals.

Our approach based on spherical Kripke models also offers insights into question (L1a) regarding the connection between Lewis counterfactuals and Stalnaker conditionals. It becomes evident that formulas of the form $\varphi > \psi$ exhibit behavior akin to Stalnaker conditionals from **VCS**. This stems from the fact that systems of spheres in a spherical Kripke model satisfy **Centering** and **Uniqueness** from Table 1. Therefore, given our definition of a Lewis counterfactual as $\varphi \mathrel{\Box\!\!\rightarrow} \psi := \Box(\varphi > \psi)$, we can provide a new insightful answer to question (L1a): a Lewis counterfactual can be interpreted as a *modality* of the corresponding Stalnaker conditional.

Moreover, our account allows for the interaction of Stalnaker conditionals, $\varphi > \psi$, and Lewis counterfactuals, $\Box(\varphi > \psi)$, at the same object language level. This expressive power enables us to model various kinds of connections between Lewis counterfactuals and Stalnaker conditionals. For instance, if we believe that a Lewis counterfactual is assertible/true only if the corresponding Stalnaker conditional is, we could introduce a reflexivity constraint on the accessibility relation to validate $\models_{\mathbf{KVC}} \Box(\varphi > \psi) \supset (\varphi > \psi)$. Indeed, this constraint establishes that a Lewis counterfactual is true only if the corresponding Stalnaker conditionals is true. Thus, not only our framework offers an answer to question (L1a) by demonstrating the definability of Lewis counterfactuals through the corresponding Stalnaker conditionals, but it also sets the ground for examining and modeling relationships between Lewis counterfactuals (and variably strict conditionals) and Stalnaker conditionals at the same object language level. Consequently, spherical Kripke models serve as a unified semantic account in which both Stalnaker conditionals and Lewis counterfactuals can coherently interact, answering question (L1c).

Formulas such as $\Box(\varphi > \psi)$ offer an alternative way to interpret Lewis counterfactual and provide new truth conditions for them, effectively addressing question (L2). In this case, a counterfactual $\varphi \mathrel{\Box\!\!\rightarrow} \psi$, now expressed as $\Box(\varphi > \psi)$, is considered true if and only if the corresponding Stalnaker conditional is true necessarily, meaning that

$$\varphi \mathrel{\Box\!\!\rightarrow} \psi \text{ is true} \quad \Leftrightarrow \quad \text{the conditional dependence expressed by } \varphi > \psi \text{ holds } \textit{necessarily}$$

Regarding the term *necessary* involved in the new truth conditions, some clarification is warranted. The discussion in Section 2.4.4 about the understanding of the modality in $LBC_\Box$ can be transferred to the current case of **KVC**. By observing the **Centering** condition in Table 3.1, we an easily derive that $\models_{\mathbf{KVC}} \Box\varphi \supset \varphi$ (Reflexivity behavior) and $\models_{\mathbf{KVC}} \Box\varphi \supset \Box\Box\varphi$ (Transitive behavior). Hence, our modal operator $\Box$ in the language $\mathcal{L}^{\Box}_{>}$ exhibit an **S4**-like behavior for formulas in the fragment $\mathcal{L}_{\restriction^{\Box}_{>}}$. Additionally, our translation σ bears some resemblances to Gödel's translation of **IL** into the modal logic **S4**. Therefore, similar the case of $LBC_\Box$, these considerations allow us to interpret the modality $\Box$ in **KVC** as expressing a notion of *provability*. Consequently, we could have rewrite the truth conditions for our counterfactual as follows:

$$\varphi \mathrel{\Box\!\!\rightarrow} \psi \text{ is true} \quad \Leftrightarrow \quad \text{the conditional dependence expressed by } \varphi > \psi \text{ is } \textit{provable}$$

Furthermore, we observe resemblances between our spherical Kripke frames and Kripke's models for intuitionistic logic. Not only does our modality $\Box$ behave similarly to an **S4** modality, but we can also establish a *hereditary* condition for counterfactual formulas: in every spherical Kripke model $\langle W, R, \mathcal{S}, f, \vDash\rangle$, for all $\varphi \in For_{\mathcal{L}_{\restriction^{\Box}_{>}}}$, for all $w, v \in W$ such that wRv, $w \vDash \varphi$ implies $v \vDash \varphi$. Hence, we can adapt Kripke's interpretation of his models for intuitionistic logic to our context, interpreting worlds in a spherical Kripke model as evidential situations and the accessibility relation as an epistemic relation between points in time. On the other hand, the hereditary condition implies that when transitioning from an actual evidential situation w to another, we do not lose our existing counterfactual information and non-conditional pieces of information. In other words, when transitioning to a world w to another accessible world world v, truth of classical formulas (non-conditional information) and of counterfactual formulas such as $\Box(\varphi > \psi)$ is preserved from w to v. Formulas true in a world w may be interpreted as expressing our available evidence at w. Hence, for a counterfactual $\Box(\varphi > \psi)$ to be true it is required that the information encoded by $\varphi > \psi$ is available at all the accessible evidential situations. Accordingly, the truth conditions for our counterfactuals may be reformulate as:

$$\varphi \mathrel{\Box\!\!\rightarrow} \psi \text{ is true} \quad \Leftrightarrow \quad \text{the corresponding Stalnaker conditional } \varphi > \psi \text{ is } \textit{certain} \text{ or } \textit{must hold} \text{ given our available evidence}$$

Alternatively, the two different accounts can be combined to obtain

$\varphi \Box\!\!\rightarrow \psi$ is true $\Leftrightarrow$ the corresponding Stalnaker conditional $\varphi > \psi$ *must be provable* given our available evidence

As mentioned earlier, our framework based on spherical Kripke models overcomes the limitations of the BACs account. Specifically, it interprets all counterfactuals in Lewis language, including nested counterfactual formulas with impossible antecedents, and faithfully translates Lewis' logic **VC** into $\models_{\mathbf{KVC}}$ logical consequence. Combining the observations made in this section with those presented in Section 2.1.3, we can conclude that we have developed a semantic account that: (i) is reductionist, as Lewis counterfactuals can be defined using the Stalnaker (or Adams) conditional equipped with a model operator; (ii) provides an alternative interpretation of the truth conditions of Lewis counterfactuals; (iii) allows for an investigation of the relationship between Lewis counterfactuals and Stalnaker conditional (or Adams conditionals) at the same object language level; (iv) serves as a unified account, analyzing Stalnaker (or Adams) conditionals and Lewis variably strict conditionals within the same semantic framework. This effectively answers all the questions (L1), (L1a), (L1b), (L1c), and (L2) posed in the introduction.

On the technical side, one one may wonder whether it is possible to axiomatize the logic associated with the logical consequence(s) $\models_{\mathbf{KV}\mathfrak{C}}$. The answer is affirmative, and it the next section, we introduce logical systems that are sound and complete with respect to the logical consequence(s) $\models_{\mathbf{KV}\mathfrak{C}}$. Interestingly, the resulting systems exhibit certain peculiar properties that connect them with the recently studied family of *weak logics* (Nakov and Quadrellaro 2022).

From a more philosophical standpoint, one significant aspect that remains absent is a comprehensive conceptual framework for evaluating and analyzing the consequences of our responses to the questions (L1), (L1a), (L1b), (L1c), and (L2). We have extensively argued that our findings present a novel comprehension of the truth conditions and logical processes underpinning Lewis counterfactuals. Furthermore, our framework exhibits greater expressiveness compared to the conventional Lewis' account, as it allows counterfactuals to interact with Stalnaker (or Adams) conditionals. For simplicity, let us focus on the framework of spherical Kripke models and the interpretation of Lewis counterfactuals using Stalnaker conditionals. However, the subsequent considerations are equally applicable to the case of the modal BACs.

From a broader perspective, one might question whether our results concerning the logic and truth conditions of Lewis counterfactuals indicate any flaws in Lewis' original construction. It could be argued that, since the truth conditions of Lewis counterfactuals can also be expressed in a more nuanced manner by invoking a modality and a Stalnaker conditional, Lewis' own understanding of his counterfactuals might have been only partial or unable to capture a deeper underlying mechanism characterizing their truth conditions. Notably, in our spherical Kripke models, we *define* Lewis' selection function, which plays a pivotal role in evaluating his counterfactuals. Thus, it becomes tempting to suggest that Lewis' truth conditions were only partial, whereas our approach reveals a more profound mechanism hidden within Lewis' truth conditions.

Conversely, one might question whether our account is inherently preferable over Lewis' or Stalnaker's, especially considering that they ultimately yield the same logic for counterfactual formulas.

Addressing these inquiries necessitates a conceptual framework that would allow us to assess and accurately evaluate the scope and implications of our results. One potential criterion for evaluation could be Van Fraassen's hidden variables theory (1974), as mentioned in the introduction. To briefly summarize, van Fraassen posits that a successful hidden variable theory $\langle T, \lambda \rangle$ should retain the correct predictions of the original theory T while offering a more comprehensive interpretation. Specifically, the element λ represents an additional parameter implicitly assumed in the original theory, which becomes explicit in the new one. In Van Fraassen's case (1974), the original theory was Lewis' logic **VC**, and the hidden variable theory was a supervaluationist extension of Stalnaker's logic, employing a supervaluationist operator $\top$. Van Fraaseen demonstrated that an expression of the form $\top(\varphi > \psi)$, where $>$ represents a Stalnaker conditional, exhibits the same logical properties as Lewis counterfactuals from **VC**.

Similarly, our theory **KVC** can be perceived as the hidden variable counterpart of Lewis' logic **VC**. In our case, the additional explanatory parameter λ is a modal operator $\Box$ that, when combined with the Stalnaker conditional, enables us to define the corresponding Lewis counterfactual and reproduce its logical behavior. In fact, through the translation σ and Theorem 3.4, our theory **KVC** replicates Lewis' inferences while offering a more profound explanation of the truth conditions of Lewis counterfactuals through the use of a modality. A successful hidden variable theory, according to van Fraassen, must not only reproduce the original theory's correct predictions but also offer superior explanatory power. Otherwise, it would fall into what van Fraassen terms "noxious metaphysics". In our case, if our theory merely reproduced Lewis' logic without introducing new semantic machinery, it could be perceived as a mere notational variant of Lewis' account, rather than making a useful addition.

However, our theory goes beyond both Stalnaker's and Lewis' logics in concrete ways. Firstly, it provides a unified account of both Stalnaker and

Lewis conditionals. Moreover, it can address specific linguistic situations that remain unexplained by either Lewis' or Stalnaker's account alone. For example, consider the question posed in the introduction:

(1) Would he get life, if he were caught?

Van Fraassen highlights that this question can be interpreted in two different ways:

(3) Is it certain (necessarily, really true) that he would get life if he were caught?

(4) Would he get a life if he were caught, or would he not get life if he were caught?

Stalnaker's theory cannot account for the first understanding, as it validates conditional excluded middle $\models_{\mathbf{VCS}} (\varphi \mathrel{\Box\!\!\rightarrow} \psi) \vee (\varphi \mathrel{\Box\!\!\rightarrow} \neg\psi)$. On the other hand, Lewis' theory cannot account for the second understanding, as it invalidates conditional excluded middle. Consequently, adhering to either of these theories would leave a gap in our explanation of everyday linguistic situations. In contrast, our theory explicitly addresses the difference between the two readings, thus filling the explanatory gap. The first reading is accounted for by making the "certain"/ "necessary" modality explicit, such that the question is formalized as asking if $\Box(\varphi > \psi)$ is true. The second reading is formalized as asking if $\varphi > \psi$ is true.

Stalnaker's theory or Lewis' theory alone cannot account for the difference between the two conditionals:

(5) If Oswald didn't kill Kennedy, someone else did

(6) If Oswald hadn't killed Kennedy, someone else would have

In contrast, our theory would formalize the first sentence as $\Box(\varphi > \psi)$ and the second as $\varphi > \psi$, or $(\psi \mid \varphi)$, thereby distinguish between the two conditionals. Therefore, based on van Fraassen's criteria, it appears that our **KVC** can be regarded as a successful hidden variable theory.

However, van Fraassen's insights, along with our results, can be seen as exemplifying a more general phenomenon in logic. In fact, an alternative and more general framework for evaluation may lie in the recently introduced conceptual distinctions between *implicit* and *explicit* stances in logic, as proposed by van Benthem (2018). While the conceptual framework of *implicit vs. explicit* is still in its developmental stages, it offers a valuable filter through which we can assess technical results connecting different logical systems. Although the conceptual methodology behind *implicit vs. explicit* has not been fully systematized, and the distinction itself remains informal, we do not necessarily require a comprehensive method to assess our results at this stage. For now, having some conceptual guidance is sufficient to adequately assess our findings, and the *implicit vs. explicit* framework can serve this purpose. We will now explore how this framework applies to our work.

Van Benthem's (2018) distinction between *implicit* and *explicit* stances in logic becomes evident through a concrete example, most notably Gödel's translation of intuitionistic logic **IL** into the modal logic **S4**. When classical *CPL* logic serves as the backdrop, **IL** is considered an *implicit* stance because it modifies or enriches our understanding of the meaning of the old logical constants or the notion of valid consequence (van Benthem 2018, p.572). For instance, **IL** provides a new semantics for the standard connective, altering their inferential behavior. A prime example is negation. In classical logic, the semantics of negation aligns with non-truth, leading to the principle of excluded middle $\varphi \vee \neg\psi$ being valid in *CPL*. However, in intuitionistic logic, the semantics of negation relates to non-provability or non-knowability, where $\neg\varphi$ is (intuitively) true when φ is not provable or not known yet. Hence **IL** doesn't introduce new linguistic devices but rather reinterprets the meaning of classical logical symbols, enriching their understanding by incorporating notions of provability of knowability. However, the provability/knowability notions are not transparent: they are *implicit* within the new intuitionistic semantics for logical connectives.

On the other hand, the modal logic **S4** can be seen as an *explicit* style of analysis concerning the notions of knowability or provability. **S4** doesn't alter the meaning or inferential behavior of the classical language; all classical inferences remain valid in **S4**, preserving the truth conditions of classical constants. Instead, **S4** expands the classical connectives with a new modal operator $\Box$, which expresses an epistemic or provability modality. Hence, the notions of knowability/provability are *explicitly* encoded within **S4** by adding new connective to express these notions, while the meaning of classical constants remain unchanged. The *implicit vs explicit* distinction between **S4** and **IL** is validated by Gödel's translation of **IL** into **S4**, demonstrating that the connectives in intuitionistic logic can be interpreted in terms of classical connectives plus an **S4** modality. This translation serves as a technical tool to uncover the the *implicit* nature of the connectives in **IL** by explaining them *explicitly* within **S4**. As van Benthem acknowledges, this translation facilitates a resounding transfer, allowing everything an intuitionist says or infers to be understood by a classical modal logician (van Benthem 2018, p.578). While the possibility of a reverse translation exists, van Benthem raises the question of whether **IL** and **S4** are merely the same system in different guises due to their faithful mutual embeddings. Addressing this question is delicate, as the mutual translation does not imply that the two systems are equivalent in all relevant aspects. In his work, van Benthem provides additional classical and new examples of the *implicit* and *explicit* stances and their mutual translations. Notably, one interesting result he proves is the translation of truthmaker semantics (Fine 2017) into modal logic.

Van Benthem's conceptual framework serves as a powerful tool for analyzing our research outcomes. Specifically, we can apply it to better comprehend the implications of both Theorem 3.4 and Theorem 2.6 which exemplify an *implicit vs explicit* relationship. To contextualize this, we posit as our backdrop a classical language equipped with a conditional operator $>$ and Stalnaker's logic (1968) (which coincides with the system **VCS**).

In Lewis' account of counterfactuals, an *implicit* style is adopted: he modifies the meaning of the conditional, $>$, (written $\Box\!\!\rightarrow$ in his account) and its inferential behavior to express a notion of counterfactual dependence. This shift alters the logical behavior of $\Box\!\!\rightarrow$, as the principle of conditional excluded middle $(\varphi \Box\!\!\rightarrow \neg\psi) \vee (\varphi \Box\!\!\rightarrow \psi)$ is valid in Stalnaker's logic, but not in Lewis' **VC**. In contrast, our system **KVC** can be considered as the *explicit* counterpart of Lewis' **VC**. We enrich Stalnaker language by introducing a new modal operator $\Box$ while keeping the meaning of classical constants and the conditional $>$ intact. This enrichment enables us to explicitly express counterfactual dependencies as $\Box(\varphi > \psi)$. By doing so, we expand the examples of *implicit vs explicit* stances proposed by van Benthem. It is worth noting that most of van Benthem's examples assume a classical language and classical logic in the background. Our approach, however, adopts Stalnaker's logic and language as the platform for comparison, representing an extension of van Benthem's conceptual framework. In fact, Lewis' account implicitly modifies the meaning of the conditional connective to convey counterfactual dependencies, while our **KVC** account explicitly analyzes counterfactual dependencies by extending Stalnaker's logic and its language. Additionally, our translation σ employed in Theorem 3.4 serves as the technical tool to establish the purported connection between Lewis' implicit approach and our **KVC** explicit account. Analogous to Gödel's translation which facilitates a transfer from **IL** to **S4**, our translation σ facilitates a transfer from **VC** (and all variably strict conditional logics) to **KVC**, thus allowing a Stalnakerian to comprehend everything Lewisian states, using a modal operator. Hence, our research results fall within the broader context of examining implicit versus explicit stances in logic.

However, this *implicit vs explicit* framework goes beyond being a philosophical category; it also holds technical and conceptual significance. Van Benthem contends that all the examples of translations from an implicit to an explicit system achieve both technical and philosophical outcomes. From a technical standpoint, these translations enable the examination of meta-logical properties of an implicit system through the lens of the corresponding explicit system. Consequently, certain properties of the latter can be easily transferred to the former through the translation process. On the philosophical side, such translations unveil new possibilities for interpreting the implicit system, shedding light on its essential features. In our study, the σ-translation played a pivotal role in revealing the concealed mechanics behind Lewis' evaluation of a counterfactual and facilitated a reexamination of its truth conditions. However, it is crucial to acknowledge that the full potential of our translation has not been entirely explored. For instance, our characterization theorem, Theorem 3.4, can be extended to encompass the entire class of variably strict conditional logics. Although we focused our philosophical considerations on counterfactuals, variably strict conditional logics express various other conditional dependencies that remain unexplored, and these can likewise be analogously reduced to a modality of a Stalnaker conditional. Moreover, from a technical perspective, this translation, by connecting Lewis conditionals with standard modal operators, may lead to the discovery of new meta-logical properties of Lewis' variably strict conditional logics which have not yet been proven. Some of these properties may naturally follow from their characterization in terms of a normal modality. Lastly, in the upcoming and final chapter, we will demonstrate how this translation contributes to resolving a longstanding problem in philosophy—the characterization of the probability expressed by a counterfactual. Specifically, the interpretation of Lewis counterfactuals in terms of a normal modal operator has been instrumental in finding an informative characterization of the associated probability. This significant advancement was made possible through our translation, which sheds new lights on the understanding of counterfactuals and their underlying probabilities.

In conclusion, the *implicit vs explicit* distinction provides a valuable conceptual framework for assessing our findings. Through these lens, our results present a compelling example of this dichotomy within the domain of conditional logic. Specifically, they reveal that Lewis' logic of counterfactuals can be viewed as an implicit approach to understanding counterfactual dependencies, while our newly introduced framework of **KVC** serves as its explicit counterpart. Broadening the perspective, as expressed by van Benthem, recognizing the existence of this *implicit vs explicit* contrast offers new avenues for exploration and leads to a deeper comprehension of logical coherence, both in various systems and in our methodological approaches. Furthermore, being aware of this contrast, particularly in our case between Lewi' counterfactual logic and our own logic **KVC**, carries profound philosophical implications:

> it undercuts sweeping ideological views that are tacitly based on taking just one design option while ignoring others. (van Benthem 2018, p.599)

Overall, the implicit vs explicit distinction not only enriches our understanding of conditional logics, but also encourages a more inclusive and thoughtful approach to its philosophical and logical analysis.

3.2.4 Logics

We have successfully demonstrated that Lewis counterfactuals (including variably strict conditionals) can be interpreted within an expanded Stalnakerian language, taking into account spherical Kripke models. Specifically, our findings have established an equivalence between any Lewis' variably strict conditional logic $\vdash_{\mathbf{V}\mathfrak{C}}$ and the logical consequence defined over the corresponding class $\mathbf{V}\mathfrak{C}$ of spherical Kripke models, with the aid of the translation σ. As previously observed, the connective $>$ involved in our definition of a counterfactual behaves logically akin to a Stalnaker conditional. Additionally, we have shed new lights on potential interpretations of the modal operator $\Box$. However, the exploration of logical consequence $\models_{\mathbf{KV}\mathfrak{C}}$ is still pending. We are aware that it encompasses Stalnaker's conditional logic, (**VCS**), Lewis' conditional logic $\vdash_{\mathbf{V}\mathfrak{C}}$, and the normal modal logic **K**. But the full extent of their interactions remains unspecified. To pursue a more comprehensive analysis of $\models_{\mathbf{KV}\mathfrak{C}}$, it is essential to ascertain whether an axiom system exists that is sound and complete with the respect to $\models_{\mathbf{KV}\mathfrak{C}}$.

In what follows, we introduce new systems of axioms and rules, giving rise to what we refer to as *modal variably strict conditional logics*, or simply **KV**-logics.

Definition 3.9. **KV** *is the logic induced by the following system of axiom schemata and rules in the language* $\mathcal{L}^{\Box}_{>}$*: (unless specified, the axiom schemata are meant to range over all formulas in the language* $\mathcal{L}^{\Box}_{>}$*)*

- ***Axioms***

  (S1) $\varphi > \varphi$

  (S2) $((\varphi > \psi) \wedge (\psi > \varphi)) \supset ((\varphi > \delta) \leftrightarrow (\psi > \delta))$

  (S3) $((\varphi \vee \psi) > \varphi) \vee ((\varphi \vee \psi) > \psi) \vee (((\varphi \vee \psi) > \delta)) \leftrightarrow ((\varphi > \delta) \wedge (\psi > \delta)))$

  (S4) $(\varphi > \psi) \supset (\varphi \supset \psi)$

  (S5) $(\varphi \wedge \psi) \supset (\varphi > \psi)$

  (S6) $(\varphi > \psi) \vee (\varphi > \neg\psi)$

  (S7) $(\varphi \gg \psi) \leftrightarrow \neg(\varphi > \neg\psi)$

(S8) $(\blacksquare\varphi) \leftrightarrow (\neg\varphi > \varphi)$

(S9) $(\blacklozenge\varphi) \leftrightarrow (\neg\blacksquare\neg\varphi)$

(S10) $(\varphi \sqsubseteq \psi) \leftrightarrow ((\varphi \vee \psi) \gg (\varphi \vee \psi)) \supset ((\varphi \vee \psi) \gg \varphi))$

(K1) $\Box(\varphi \supset \psi) \supset (\Box\varphi \supset \Box\psi)$

(K2) $\neg\Box\neg\varphi \leftrightarrow \Diamond\varphi$

(KV1) *for* $\varphi, \psi, \delta \in For_{\mathcal{L}^{\Box}_{>}}$,
$\Box((\varphi \vee \psi) > \varphi) \vee \Box((\varphi \vee \psi) > \psi) \vee (\Box((\varphi \vee \psi) > \delta)) \leftrightarrow \Box((\varphi > \delta) \wedge (\psi > \delta)))$

- ***Rules***

(RS1) *Modus Ponens*

(RS2) $\vdash \varphi$ *when* φ *is a classical tautology*

(RS3) $\varphi_1, \varphi_2, \ldots, \varphi_n \vdash \varphi$ *when* $(\varphi_1 \wedge \cdots \wedge \varphi_n) \supset \varphi$ *is a classical tautology*

(RS4) *if* $\vdash \varphi$, *then* $\vdash \psi > \varphi$

(RS5) *if* $\vdash (\varphi_1 \wedge \cdots \wedge \varphi_n) \supset \varphi$ *then* $\vdash ((\psi > \varphi_1) \wedge \cdots \wedge (\psi > \varphi_n)) \supset (\psi > \varphi)$

(RS6) *substitution of interderivable formulas*

(RK1) *if* $\vdash \varphi$, *then* $\vdash \Box\varphi$

For an axiom or a family of axioms (possibly empty) $\mathfrak{C}$ *among those in Table 3.1 (i.e.* $\{\mathbf{N}, \mathbf{T}, \mathbf{W}, \mathbf{C}, \mathbf{S}, \mathbf{A}, \mathbf{U}, \}$*),* $\mathbf{KV}\mathfrak{C}$ *is the logic induced by the system obtained by extending* **KV** *with the axioms in* $\mathfrak{C}$.

For $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}^{\Box}_{>}}$, *a derivation from* Γ *to* φ *in* $\mathbf{KV}\mathfrak{C}$ *is a finite sequence of formulas that ends with* φ *such that each formula in the sequence instantiate the axioms in* $\mathbf{KV}\mathfrak{C}$, *or belongs to* Γ *or it is obtained by applications of he rules in* $\mathbf{KV}\mathfrak{C}$.

For $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}^{\Box}_{>}}$, $\Gamma \vdash_{\mathbf{V}\mathfrak{C}} \varphi$ *means that there is a derivation from formulas in* Γ *to* φ.

From the definition of derivation, it is straightforward to see that for $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}^{\Box}_{>}}$, *for an axiom or a family of axioms (possibly empty)* $\mathfrak{C}$ *among those in Table 3.1 (i.e.* $\{\mathbf{N}, \mathbf{T}, \mathbf{W}, \mathbf{C}, \mathbf{S}, \mathbf{A}, \mathbf{U}, \}$*)*

$$\Gamma \vdash_{\mathbf{KV}\mathfrak{C}} \varphi \Leftrightarrow \Delta \vdash_{\mathbf{KV}\mathfrak{C}} \varphi \text{ for some finite } \Delta \subseteq \Gamma$$

Let us examine the logic(s) just introduced. The axioms $(S1) - -(S10)$ are essentially borrowed from Stalnaker's logic **VCS** , along with the rules $(RS1) - -(RS6)$. The axioms $(K1) - (K2)$, along with the rule $(KR1)$, are borrowed from the normal modal logic **K** (see Blackburn, de Rijke, and Venema 2001). The only axiom that can be considered "new" is $(KV1)$, which corresponds to the **Sphericality** condition on the function f over spherical

Kripke models. Roughly speaking, similar to the model-theoretic case where spherical Kripke models results from merging together a Kripke frame and a spherical Stalnakerian model while adding a **Sphericality** condition, the basic logic **KV** arising from combining the normal modal logic **K** and Stalnaker's logic **VCS**, supplemented by the new characteristic axiom $(KV1)$. We have considered all the possible axiomatic extension of **KV** by adding the axioms in Table 3.1. In principle, it is also feasible to consider further axiomatic extensions of **KV**, by including normal modal logic axioms, such as $\Box\varphi \supset \varphi$ for all $\varphi \in For_{\mathcal{L}^{\Box}_{>}}$. The interplay between normal modal logic axioms and the axioms in Table 3.1 enhances the expressive power of these newly introduced logics. For instance, it is evident that if we extend **KV** with the axiom $\mathbf{T} : \Box\varphi \rightarrow \varphi$, and the axiom $(\varphi \wedge \psi) \supset (\varphi \mathrel{\Box\!\!\rightarrow} \psi)$ (for all $\varphi, \psi \in For_{\mathcal{L}^{\Box}_{\restriction >}}$), the resulting logic coincides with Lewis counterfactual logic **VC**, modulo the translation σ.

We will now show that $\mathbf{KV}\mathfrak{C}$ logics are sound and complete with respect to the corresponding logical consequence $\models_{\mathbf{KV}\mathfrak{C}}$.

Theorem 3.5. *Let $\mathfrak{C}$ be an axiom/condition or a family of axioms/conditions (possibly empty) among those in Table 3.1, (i.e. $\{\mathbf{N}, \mathbf{T}, \mathbf{W}, \mathbf{C}, \mathbf{S}, \mathbf{A}, \mathbf{U}, \}$), then, for all $\Gamma \cup \{\varphi\} \subseteq For_{\mathcal{L}^{\Box}_{>}}$, the following holds:*

$$\Gamma \vdash_{\mathbf{KV}\mathfrak{C}} \varphi \Leftrightarrow \Gamma \models_{\mathbf{KV}\mathfrak{C}} \varphi$$

Proof. The proof is based on merging together the standard completeness proof for normal modal logics and variably strict conditionals logic, by using our Lemma 3.3. We won't provide all the details, bit we will include a detailed idea of the proof in Appendix C.2 $\square$

An extensive investigation of the properties of these logics goes beyond the scope of the present thesis. Our primary objective was to address questions (L1), (L1a), (L1b), and (L1c) posed in the introduction. Nevertheless, a notable characteristic of these **KV**-logics deserves attention. Specifically, inferences within these logics are not closed under uniform substitution. For instance, we can easily observe that:

$$\vdash_{\mathbf{KV}} \Box((p \vee p) > p) \vee \Box((p \vee q) > q) \vee (\Box((p \vee q) > c)) \leftrightarrow \Box((p > c) \wedge (q > c)))$$

is a theorem of **KV**, by axiom (KV1). However, its substitutional instance, where p is replaced by $p > q$ and q by $q > p$

$$\nvdash_{\mathbf{KV}} \Box(((p > q) \vee (q > p)) > (p > q)) \vee \Box(((p > q) \vee (q > p) > (q > p)) \vee (\Box(((p > q) \vee (q > p)) > c)) \leftrightarrow \Box(((p > q) > c) \wedge ((q > p) > c)))$$

is *not* valid in with respect to the logical consequence $\models_{\mathbf{KV}}$, and thus it is not a theorem. This raises a natural question regarding whether **KVC** can truly be considered as *logics*.

Addressing the philosophical issue of whether **KV**-logics can be properly classified as logics is beyond the scope of this thesis. However, it is crucial to note that there exists a diverse range of "logics" that are not closed under uniform substitution, yet they offer intriguing applications and meta-logical properties. A prominent example is inquisitive logic (Ciardelli 2022), which finds applications to questions and inquisitive reasoning. Within this realm, there is a special family of "logics", including inquisitive logics, known as *weak logics*, which exhibit an interesting behavior and can be effectively modeled in algebraic terms. Weak logics have been introduced and studied only very recently by Nakov and Quadrellaro (2022); specifically, weak logics are those logics that maintain inferences under atomic uniform substitutions. More formally:

Definition 3.10 (Nakov and Quadrellaro 2022)**.** *A weak logic is a finitary consequence relation* $\vdash$ *such that for all atomic substitutions* s, $\Gamma \vdash \varphi$ *implies* $s[\Gamma] \vdash s(\varphi)$ *where* $s(\varphi)$ *is the result of substituting all the occurrences of a variable* p *in* φ *with another (possibly the same) variable* q*, and* $s[\Gamma] = \{s(\gamma) \mid \gamma \in \Gamma\}$

Evidently, our **KV**-logics fall under the category of weak logics as per the given definitions. Nakov and Quadrellaro (2022) introduced a notion of algebraizability for weak logics, along with a method to develop an algebraic treatment for them. Consequently, our **KV** can also be analyzed from an algebraic perspective. In the future, we aim to delve more deeply into the properties of **KV**-logics by connecting them to the more general theory of weak logics by Nakov and Quadrellaro (2022). For now, our primary objective is to show that **KV**-logics are not just *ad hoc* systems, but are integrated into a novel and expanding logical framework.

3.3 Conclusions

In summary, our research has demonstrated how Lewis counterfactual conditionals can be defined using the connectives in the $\mathcal{L}^{\Box}_{>}$ within the framework of spherical Kripke models. Specifically, a counterfactual can be expressed as $\varphi \mathrel{\Box\!\!\rightarrow}^{G} \psi := \Box(\varphi > \psi)$, where $>$ corresponds to a Stalnaker conditional, and $\Box$ is a normal modal operator satisfying the axiom (KV1) and **Centering**. The counterfactual formulas $\Box(\varphi > \psi)$ exhibit a logical behavior analogous to Lewis counterfactuals. Notably, Lewis' logic of counterfactuals **VC** can be embedded into our corresponding **KVC** logical consequence through our translation σ, and the same result extends to the entire family of Lewis' variably strict conditional logics. Our framework not only replicates Lewis' logics, but also provides a unified treatment at the object language level for both Stalnaker conditionals and Lewis counterfactuals. Based on these logical results, we have put forth a new interpretation of the truth conditions for Lewis counterfactuals. In particular, a Lewis counterfactual $\varphi \mathrel{\Box\!\!\rightarrow} \psi$ can be interpreted as the corresponding formulas $\Box(\varphi > \psi)$, thus being understood as a *modality* of the corresponding Stalnaker conditional. Moreover, we have offered plausible insights into the modality expressed by $\Box$. In analogy with Gödel's translation of intuitionistic logic and Kripke semantics for intuitionistic logic, our framework shows how the modality can be interpreted as a *provability* or *epistemic* modality. This sheds new lights on the underlying meaning and significance of the modal operator within our framework. In addition to axiomatizing the logic induced by the $\models_{\mathbf{KV}\mathfrak{C}}$ logical consequence over our spherical Kripke models, we have observed that these logics differ from Tarskian logical relations by not preserving valid inferences under uniform substitution. However, they fall under the newly introduced category of weak logics, which includes intriguing examples like inquisitive logic. Furthermore, we have evaluated our framework and the associated results from a broader philosophical perspective. According to van Fraassen's criteria, our framework successfully qualifies as a hidden variable theory. It reproduces the same logic induced by Lewis' framework while offering a more insightful interpretation of it in terms of a modality and a Stalnaker conditional. On the other hand, following van Benthem's analysis, our **KVC** account can be regarded as an *explicit* analysis of counterfactual dependencies, whereas Lewis' theory represent its *implicit* counterpart. The dichotomy between implicit and explicit stances in logic has been developed as a conceptual filter to make sense of and interpret translations between different logical systems, such as the one between our theory and Lewis' theory. Moving forward, in the last chapter of this thesis, we delve into a probabilistic analysis of Lewis counterfactuals.

Chapter 4

Probability of Counterfactuals (and Variably Strict Conditionals)

The present chapter focuses on the characterization of the probability of (the proposition expressed by) a Lewis counterfactual, addressing question (P1) mentioned in the introduction.

4.1 Background

In this section, we recap the basic notions needed to provide our characterization result for the probability of a counterfactual.

4.1.1 Probability

Definition 4.1. *Given a finite Boolean algebra* $\mathbf{A} = \langle A, \wedge, \vee, \neg, \top, \bot \rangle$*, a probability* P *over* $\mathbf{A}$ *is a function* $P : \mathbf{A} \to [0, 1]$ *from* $\mathbf{A}$ *to the real interval* $[0, 1]$ *satisfying the following conditions:*

1. $P(\top) = 1$
2. *if* $a \wedge b = \bot$*, then* $P(a \vee b) = P(a) + P(b)$

We say that P *is positive when for all* $\alpha \in at(\mathbf{A})$, $P(\alpha) > 0$.

Furthermore, given a distribution P over $at(\mathbf{A})$, $P : at(\mathbf{A}) \to [0, 1]$, P can be extended to a probability $P : \mathbf{A} \to [0, 1]$ over $\mathbf{A}$ as follows:

$$P(a) = \sum_{\alpha \leq a} P(\alpha)$$

As a consequence, we obtain the following:

Lemma 4.1. *Given a non empty set of possible worlds W, consider the Boolean algebra $\langle \wp(W), \cap, \cup, ^{-}, \emptyset, W\rangle$. A distribution P over the elements of W, $P : W \longrightarrow [0,1]$, can be extended to a probability function over $\wp(W)$ as follows: for $X \subseteq W$*

$$P(X) = \sum_{w \in X} P(w)$$

Since a proposition is a subset of $X \subseteq W$, $P(X)$ is the probability of the proposition X.

The above observations set the ground for the assignment of a probability to the proposition expressed by a counterfactual. First, we introduce some useful terminology:

Notation 4.1.

> *Given a functional (or, equivalently, a spherical) Lewisian model $\mathcal{M} = \langle W, f, \vDash\rangle$, we use the following terminology:*
>
> - *for a formula $\varphi \in For_{\mathcal{L}_{\Box\!\rightarrow}}$, $[\varphi] = \{w \in W \mid w \vDash \varphi\}$ is the proposition expressed by φ, i.e. the set of worlds where φ is true*

It is straightforward to observe that a finite functional Lewisian model induces a Boolean algebra, hence a probability distribution can be defined on it according to Definition 4.1.

Remark 4.1. *Given a finite functional Lewisian model $\mathcal{M} = \langle W, f, \vDash\rangle$, the structure $\langle \wp(W), \cap, \cup, ^{-}, W, \emptyset\rangle$ is a Boolean algebra. Moreover, observe that for all $\varphi \in For_{\mathcal{L}_{\Box\!\rightarrow}}$, $[\varphi] \in \wp(W)$. Namely, the proposition expressed by φ, i.e. $[\varphi] = \{w \in W \mid w \vDash \varphi\}$, is an element of the Boolean algebra $\langle \wp(W), \cap, \cup, ^{-}, W, \emptyset\rangle$ induced by $\mathcal{M}$. In particular, for $\varphi, \psi \in For_{\mathcal{L}_{\Box\!\rightarrow}}$, $[\varphi \Box\!\!\rightarrow \psi] \in \wp(W)$*

From the above remark and Definition 4.1, we can easily see how to compute the probability of (the proposition expressed by a) Lewis counterfactual:

Remark 4.2. *Given a functional Lewisian model $\mathcal{M} = \langle W, f, \vDash\rangle$ satisfying* **Centering** *and a probability distribution $P : W \longrightarrow [0,1]$ over W, for $\varphi \in For_{\mathcal{L}_{\Box\!\rightarrow}}$ the probability of (the proposition expressed by) φ is:*

$$P(\varphi) = P([\varphi]) = \sum_{w \vDash \varphi} P(w)$$

and in particular

$$P(\varphi \Box\!\!\rightarrow \psi) = \sum_{w \vDash \varphi \Box\!\rightarrow \psi} P(w)$$

In order to avoid verbosity, we adopt the following terminology:

Notation 4.2.

- *Given a (finite) functional Lewisian model $\mathcal{M} = \langle W, f, \vDash \rangle$ and a probability distribution $P : W \to [0,1]$, for every formula $\varphi \in For_{\mathcal{L}_{\Box\!\!\to}}$, we refer to the probability of the proposition expressed by φ as the probability of φ.*

Namely, given a finite functional (or spherical) Lewis model $\mathcal{M}$ satisfying **VC** and a probability distribution over the possible worlds in $\mathcal{M}$, the probability of a counterfactual $\varphi \Box\!\!\to \psi$ is available and easily computed as the cumulative sum of the probabilities of the worlds at which $\varphi \Box\!\!\to \psi$ is true. However, a natural question arises concerning how we should interpret the probability of a counterfactual $P(\varphi \Box\!\!\to \psi)$. On an intuitive level, a probability distribution P over a set of worlds W provides information on how likely it is that a certain world W is our actual world. For instance, for $w \in W$, $P(w) = 0.5$ would intuitive mean that there is fifty percent of chances that w is our actual world. Under this intuitive understanding, then, for $\varphi \in For_{\mathcal{L}_{\Box\!\!\to}}$, $P(\varphi)$ would amount to the chance that φ holds at the actual world, that is, how probable it is that φ is true. However, simply having

$$P(\varphi \Box\!\!\to \psi) = \sum_{w \vDash \varphi \Box\!\!\to \psi} P(w)$$

and its intuitive understanding, i.e. the probability that $\varphi \Box\!\!\to \psi$ is true, does not provide any substantial information on the *nature* of the probability of $\varphi \Box\!\!\to \psi$. For instance, in principle, we cannot know what is the relation between $P(\varphi \Box\!\!\to \psi)$ and the corresponding conditional probability $P(\psi \mid \varphi)$ unless we find a more informative characterization of $P(\varphi \Box\!\!\to \psi)$. Let us consider an analogous and much easier example. $P(\varphi \vee \psi)$ is the probability that either φ or ψ are true. However, we also know that $P(\varphi \vee \psi) = P(\varphi) + P(\psi) - P(\varphi \wedge \psi)$, namely $P(\varphi \vee \psi)$ amounts to the sum of the probabilities of φ and ψ, minus the probability of their conjunction. Hence, for example, for $\varphi \vee \psi$ to have a fifty percent chance of being true, it is required that the sum of $P(\varphi)$ and $P(\psi)$ minus $P(\varphi \wedge \psi)$ must be 0.5. In the case of a counterfactual the question becomes: what is required for $\varphi \Box\!\!\to \psi$ to have x of chances of being true? What does it mean to ask how probable it is that $\varphi \Box\!\!\to \psi$ is true?

4.1.2 The Probability of a Counterfactual is not an Imaged Probability

As we anticipated in the introduction, in the relevant literature on the topic, there are various accounts of the probability of a counterfactual. Many of them rely on the idea that the probability of a counterfactual $P(A \boxright B)$ coincides with a *counterfactual probability*, i.e. $P(A \boxright B)$ coincides with the probability of *A under the counterfactual assumption that A holds*. We will focus on one main proposal in this realm, arguing that it fails to fully characterize the probability of the proposition expressed by a Lewis counterfactual.

Generalized Imaging

Generalized imaging is a specific type of updating procedure; it has also been proposed as an alternative to Bayesian conditionalization in the context of Causal Decision Theory (Joyce 1999). Generalized imaging has been originally introduced by Gärdenfors (1982) as a generalization of Lewis' imaging, indeed the latter can be presented as particular instantiation of the former.

The generalized imaging approach is employed to model an updated belief state. Let us assume that our current belief state is correctly modeled by a probability distribution P over a space of possible worlds W, and P extends to a probability function over the algebra of propositions $\wp(W)$. Assume that, at some point, we acquire new knowledge that φ is the case. How should we update our probability distribution P in order to account for this newly acquired knowledge? One possibility is Bayesian conditionalization: P becomes the new function $P(\cdot \mid \varphi)$, so that for each formula ψ, its new probability becomes $P(\psi \mid \varphi) = \frac{P(\psi \wedge \varphi)}{P(\varphi)}$. Under this new updated probability, it is straightforward to observe that $P(\varphi \mid \varphi) = 1$, meaning that φ has now become certain.

An alternative updating procedure consist in the generalized imaging (Gärdenfors 1982; Günther 2022):

Definition 4.2. *Given a functional Lewisian model $\mathcal{M} = \langle W, f, \vDash \rangle$ satisfying* **Centering**, *consider a probability distribution P over W. Moreover, assume that for all $\varphi \in For_{\mathcal{L}_{\boxright}}$ such that $\nvDash_{\mathbf{VC}} \neg\varphi$, for all $w \in W$, $f(\varphi, w) \neq \emptyset$.*

Consider a formula $\varphi \in For_{\mathcal{L}_{\boxright}}$, and a function $T_\varphi : W \times W \to [0,1]$ subject to the following constraint: for all $w, v \in W$, for all φ, if $f(\varphi, w) \neq \emptyset$, then

$$\sum_{v \in f(\varphi, w)} T_\varphi(w, v) = 1$$

Then, the imaged probability distribution P^φ is defined as follows: for all $w \in W$,

$$P^\varphi(w) = \sum_{v \in W} P(v) \times \begin{cases} T_\varphi(v,w) & \textit{if } v \in f(\varphi,w) \\ 0 & \textit{otherwise} \end{cases}$$

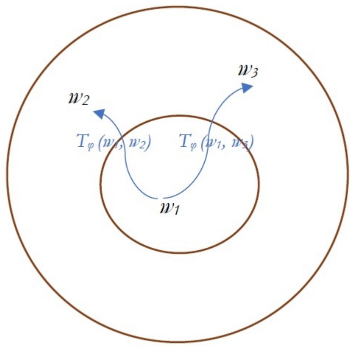
Intuitively, the above definition establishes a procedure for updating an original probability distribution P upon learning new information φ. This procedure works as follows: we update P to obtain a new distribution P^φ that incorporates the newly acquired evidence encoded by φ. In this updated distribution, each possible world w such that $w \nvDash \varphi$ (i.e. where φ is false) must *lose* its weight $P(\varphi)$, resulting in $P^\varphi(w) = 0$. Conversely, all the φ-worlds keep their original weight. However, the φ-worlds can gain additional weight under P^φ. To elaborate further, let us consider a world w such that $w \nvDash \varphi$. This world *loses* its original weight $P(w)$, in $P^\varphi(w)$. However, $P(w)$ is *redistributed* among the most similar worlds to w that satisfy φ, denoted as $f(\varphi,w)$. The transfer of weight is determined by T_φ, which encodes the proportion of the weight of w that must be transferred to each world in $f(\varphi,w)$. For example, if $f(\varphi,w)$ contains only one element v, then all the weight of $P(w)$ is transferred to v. On the other hand, if $f(\varphi,w)$ contains more than one element, then each $v \in f(\varphi,w)$ receives a fraction of the original weight of w, given by $T_\varphi(w,v)$, i.e. that fraction would be equal to $P(w) \times T_\varphi(w,v)$. Consequently, while $\neg\varphi$-worlds lose all their original weight, φ-worlds retain their original weight and may gain even more weight under P^φ.

Example of Generalized Imaging

Consider a functional Lewisian model $\mathcal{M} = \langle W, f, \vDash \rangle$ equipped with a probability distribution over W, where:

- there is $w_1 \in W$ such that $P(w_1) = 0.5$
  - $f(\varphi, w_1) = \{w_2, w_3\}$
- $T_\varphi(w_1, w_2) = \frac{1}{3}$
- $T_\varphi(w_1, w_3) = \frac{2}{3}$

This means that w_2 would gain $\frac{1}{3}$ of 0.5, which is the original weight of w_1, under the imaged P^φ, and w_3 would gain $\frac{2}{3}$ of 0.5. Then, the resulting imaging probability is depicted in the following figure:



$$P(w) = 0.5 \rightsquigarrow P^{\varphi}(w) = 0$$
$$P(w_2) = 0.2 \rightsquigarrow P^{\varphi}(w_2) = P(w_2) + \tfrac{1}{3} \times P(w)$$
$$P(w_3) = 0.3 \rightsquigarrow P^{\varphi}(w_3) = P(w_3) + \tfrac{2}{3} \times P(w)$$

Figure 4.1: A pictorial representation of generalized imaging

The following corollary easily follows from the definition of generalized imaging:

Corollary 4.1. *Given a functional Lewisian model $\mathcal{M} = \langle W, f, \vDash \rangle$ satisfying* **Centering** *and* **Uniqueness***, consider a probability distribution P over W. Moreover, assume that for all $\varphi \in For_{\mathcal{L}_{\Box\!\rightarrow}}$ such that $\not\vDash_{\mathbf{VCS}} \neg\varphi$, for all $w \in W$, $f(\varphi, w) \neq \emptyset$.*

Consider a formula $\varphi \in For_{\mathcal{L}_{\Box\!\rightarrow}}$, and a function $T_{\varphi} : W \times W \longrightarrow [0, 1]$ as in Definition 4.2. The imaged probability distribution P^{φ} over W can be characterized as follows:

$$P^{\varphi}(w) = \sum_{v \in W} P(v) \times \begin{cases} 1 & \text{if } v \in f(\varphi, w) \\ 0 & \text{otherwise} \end{cases}$$

Observe that the above corollary applies to functional models satisfying **Centering** and **Uniqueness**, namely to models of Stalnaker's logic **VCS**. Indeed, the following result has been proved by Lewis:

Theorem 4.1 (Lewis 1976)**.** *Given functional Lewisian model $\mathcal{M} = \langle W, f, \vDash \rangle$ satisfying* **Centering** *and* **Uniqueness***, consider a probability distribution P over W. Moreover, assume that for all $\varphi \in For_{\mathcal{L}_{\Box\!\rightarrow}}$ such that $\not\vDash_{\mathbf{VCS}} \neg\varphi$, for all $w \in W$, $f(\varphi, w) \neq \emptyset$. Then $P^{\varphi}(\psi) = \sum_{w \vDash \psi} P^{\varphi}(w) = P(\varphi \Box\!\rightarrow \psi)$.*

Remark 4.3. *Since $\mathcal{M}$ satisfies* **Centering** *and* **Uniqueness***, the conditional $\Box\!\rightarrow$ in $\mathcal{M}$ is essentially a Stalnaker conditional. Hence, in order to avoid confusion, we write $>$ in place of $\Box\!\rightarrow$ within a functional Lewisian model $\mathcal{M} = \langle W, f, \vDash \rangle$ satisfying* **Centering** *and* **Uniqueness***. Hence, the above result can be rewritten as:*

$$P^{\varphi}(\psi) = \sum_{w \vDash \psi} P^{\varphi}(w) = \sum_{w \vDash \varphi > \psi} P(w) = P(\varphi > \psi)$$

The above theorem, due to Lewis establishes that the probability of a Stalnaker conditional $P(\varphi > \psi)$ is the imaged probability of ψ upon φ, $P^{\varphi}(\psi)$. The probability $P^{\varphi}(\psi)$ results from updating the original P upon learning φ. Since we are in the context of models for **VCS**, the lost mass of a $\neg\varphi$-world w is entirely transferred to its closest φ-world in $f(\varphi, w)$.

So, Lewis has provided a characterization of the probability of a Stalnaker conditional in terms of imaging within the context of **VCS** models. It would seem natural if the (generalized) imaging procedure could be applied to the case of **VC** models, yielding a characterization of the probability of a Lewis counterfactual. In fact, Günther (2022) and, on different grounds, Schulz (2017) have argued that generalized imaging over **VC** functional models characterizes the probability of a Lewis counterfactual. However, this is not case. In particular, we can find a functional Lewisian **VC** model in which the probability of a Lewis counterfactual does not coincide with the corresponding imaged probability.

Remark 4.4. *The probability of a Lewis counterfactual, i.e. a conditional obeying the logic* **VC**, *is not characterized by generalized imaging as in Definition 4.2.*

Proof. Consider a spherical Lewisian model $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$ satisfying **Centering**, namely we allow for the possibility that $f(\varphi, w)$ contains more than one element. Then $\mathcal{M}$ is a model for Lewis' logic **VC**, i.e. a model for Lewis counterfactuals. In particular, we assume that $\mathcal{M}$ has the following structure:

- $W = \{w_1, w_2, w_3\}$
- $\mathcal{S}$ is such that:
  - $\mathcal{S}(w_1) = \{\{w_1\}, \{w_1, w_2, w_3\}\}$
  - $\mathcal{S}(w_2) = \{\{w_2\}\}$
  - $\mathcal{S}(w_3) = \{\{w_3\}\}$
- $w_1 \nvDash p$, $w_2 \vDash p$, $w_3 \vDash p$. Moreover; $w_3 \nvDash q$, $w_2 \vDash q$

Now, consider the functional Lewisian model induced by $\mathcal{M}$, according to Definition 0.4, i.e. $\mathcal{M}^f = \langle W, f, \vDash \rangle$. It is straightforward to see that:

- $f(p, w_1) = \{w_2, w_3\}$

Now, consider the probability distribution P over W such that:

- $P(w_1) = P(w_2) = P(w_3) = \frac{1}{3}$

and set $T_p(w_1, w_2) = 0.5$ and $T_p(w_1, w_3) = 0.5$. Then the resulting imaged distribution P^p would be:

- $P^p(w_1) = 0$
- $P^p(w_2) = \frac{1}{3} + (0.5 \times \frac{1}{3}) \approx 0.49$
- $P^p(w_3) = \frac{1}{3} + (0.5 \times \frac{1}{3}) \approx 0.49$

It is easy to see that $P^p(q) = P^p(w_2) = 0.49$ but $P(p \mathrel{\Box\!\!\rightarrow} q) = P(w_2) \approx 0.33$, namely

$$P^p(q) \neq P(p \mathrel{\Box\!\!\rightarrow} q)$$

□

The above observation establishes that generalized imaging fails to capture the probability of a Lewis counterfactual in the following sense: there are at least one Lewisian model $\mathcal{M}$ satisfying **Centering**, two formulas $\varphi, \psi \in For_{\mathcal{L}_{\Box\rightarrow}}$, a probability distribution on $\mathcal{M}$, and an imaging function T_φ such that

$$P^\varphi(\psi) \neq P(\varphi \mathrel{\Box\!\!\rightarrow} \psi)$$

However, one might consider the possibility of restricting our focus to a special kind of imaging procedures that could restore the equation between the probability of Lewis counterfactuals and imaged probability. For instance, in the model provided in the proof of Remark 4.4, we could tune T_p in such a way that $T_p(w_1, w_2) = 1$. Under this assumption, then, $P^p(q) = P(\varphi \mathrel{\Box\!\!\rightarrow} \psi)$. Hence, this suggests that the characterization of the probability of Lewis counterfactuals in terms of generalized imaging might work only for a specific class of imaging rules,. However, it is essential to notice that tuning T^p in the way described implies that the non-equality strikes back for $p \mathrel{\Box\!\!\rightarrow} (p \wedge \neg q)$, in particular, we would have that $P^p(p \wedge \neg q) \neq P(p \mathrel{\Box\!\!\rightarrow} (p \wedge \neg q))$

From the above observations, we could deduce that generalized imaging fails to fully characterize the probability of a Lewis counterfactual, disproving the claims made by Günther (2022) and Schulz (2017). In what follows, we lay the groundwork to provide a faithful characterization of the probability of a Lewis counterfactual.

4.2 A Faithful Characterization of the Probability of a Lewis Counterfactual

4.2.1 Preliminaries: Belief Functions

Dempster-Shafer Theory is a formal framework theorized by Dempster (1968) and formally developed by Shafer (1976) to model reasoning under uncertainty. The framework can be regarded as a generalization of classical Bayesian probability, as some principles of classical probability are relaxed within the Dempster-Shafer framework. To illustrate the informal ideas behind how Dempster-Shafer theory models uncertainty, consider the following example.

Suppose an assassination was committed, and there are three candidate suspects for the murder: *John*, *Mary*, and *Peter*, and we know that only one of them is guilty[1]. Hence, our sample space in this context can be modeled as the three-element Boolean algebra, where atoms/possible worlds are the three states where:

1. w_1 :*John* is guilty, *Mary* and *Peter* are innocent.
2. w_2: *Mary* is guilty, *John* and *Peter* are innocent.
3. w_3 *Peter* si guilty, *Mary* and *John* are innocent.

During the investigation, a witness of the episode truthfully affirms to have seen the murder escaping in the dark after the assassination and noticed that that person he was a *man*. However, we know that this witness is reliable 80% of the times. The question is how to model our epistemic uncertainty in this context, specifically, how to distribute our mass among the possible worlds w_1, w_1, and w_2 and the space of propositions $\wp(\{w_1, w_2, w_3\})$. According to the information provided by our witness, the proposition that *John is guilty or Peter is guilty* should be true with 80% chance. In standard probability theory, we would distribute our mass among the three possible worlds in such a way that $P(w_1) + P(w_2) + P(w_3) = 1$ and $P(\textit{John is guilty or Peter is guilty}) = P(w_1) + P(w_3) = 0.8$. Since we lack any other relevant evidence besides the information provided by the witness, standard probability theory would lead us to adopt a *principle of indifference* and distribute the mass in such a way that w_1 and w_2 are equally likely. That implies that, based on our evidence, it is equally likely that John or Peter committed the murder. However, this assumption appears strong if we consider that we have no real information about whether Peter and John are potentially equally guilty.

On the other hand, in the context of Dempster-Shafer theory, our epistemic uncertainty in the above case would be modeled by assigning the mass 0.8 to the proposition expressed by *John is guilty or Peter is guilty*, i.e. to the whole set $\{w_1, w_2\}$ rather than the single possible worlds. In this context, we are not committed to the string assumption implied by the principle of indifference in the classical setting. The next definition will clarify the mass distribution in line with the idea behind Dempster-Shafer theory:

Definition 4.3. *A mass distribution over a finite Boolean algebra* $\mathbf{A} = \langle A, \wedge, \vee, \neg, \top, \bot \rangle$ *is a function* $m : \mathbf{A} \to [0,1]$ *from* $\mathbf{A}$ *to the real interval* $[0,1]$

[1]This example is adapted from (Denoeux 2011).

such that

$$\sum_{a \in A} m(a) = 1$$

Consequently, the following lemma easily follows:

Lemma 4.2. *Given a non empty set of possible worlds W, consider the Boolean algebra* $\langle \wp(W), \cap, \cup, ^{-}, \emptyset, W \rangle$. *A function m over the elements of* $\wp(W)$, *such that* $m : \wp(W) \rightarrow [0,1]$ *and* $\sum_{X \subseteq W} P(X) = 1$, *is a mass distribution on* $\wp(W)$.

Since a proposition is a subset of $X \subseteq W$, $m(X)$ *is the mass assigned to the proposition* X.

Let us compare the above definition with Definition 2.10. While classical probability distributes the mass among the possible worlds (atoms of Boolean algebra), a mass distribution in Dempster-Shafer theory assigns masses to sets of possible worlds, i.e. propositions (elements of a Boolean algebra).

In the previous section, we have intuitively seen that, under a classical probability assignment P, the weight assigned to a possible world may be interpreted as the chance that the world is the actual one. Accordingly, the weight assigned by P to a proposition, may be regarded as the probability for that proposition to be true.

On the other hand, the weight assigned by a mass distribution m to a proposition X, denoted as $m(X)$, may be seen as the chance that the actual world falls into that proposition X, or alternatively, as the quantification of the strength of the evidence we have in favor of that proposition. For instance, in the case of the assassination example, we could have a mass distribution such that $m(\{w_1, w_3\}) = 0.8$, representing the "strength" of the evidence we have in favor of the proposition *John is guilty or Peter is guilty*. From a mass distribution, we can construct the fundamental technical tool of Dempster-Shafer theory, that is *Belief functions*.

Definition 4.4. *A belief function on a finite Boolean algebra* $\mathbf{A}$ *is a map* $Bel : \mathbf{A} \rightarrow [0,1]$ *satisfying the following properties:*

- $Bel(\top) = 1$;
- $Bel(a \vee b) \geq Bel(a) + Bel(b) - Bel(a \wedge b)$

A belief function Bel is said to be normalized if $Bel(\bot) = 0$. *Moreover, given a mass distribution* $m : \mathbf{A} \rightarrow [0,1]$, *the map* $Bel : \mathbf{A} \rightarrow [0,1]$ *defined as follows*

$$Bel(a) = \sum_{x \leq a} m(x)$$

is a belief function over $\mathbf{A}$, *and every belief function on* $\mathbf{A}$ *arises in this way.*

Consequently, we have the following characterization of Belief functions over power-sets Boolean algebras.

Lemma 4.3. *Consider a non-empty set of possible worlds W and its associated Boolean algebra $\langle \wp(W), \cap, \cup, ^-, \emptyset, W \rangle$. Consider a mass distribution $m : (\wp(W)) \rightarrow [0,1]$, the map $Bel : \wp(W) \rightarrow [0,1]$ defined as follows*

$$Bel(X) = \sum_{X \in \wp(W) : Y \subseteq X} m(Y)$$

is a belief function over $\wp(X)$.

Namely, belief functions arise from mass distributions, in an analogous way as probability functions arise from probability distributions.

We have provided an intuitive understanding of the mass distribution: $m(X)$ can be seen as a measure of the evidence we have in favor of the proposition X. Now, a natural question arises concerning how we should interpret $Bel(X)$. According to Definition 4.4, a belief function of some proposition X is:

$$Bel(X) = \sum_{X \in \wp(W) : Y \subseteq X} m(Y)$$

namely, $Bel(X)$ is the cumulative sum of all the pieces of evidence, $m(Y)$, for all proposition Y that are contained in X. Given this observation, an illuminating interpretation of $Bel(X)$ is due to Pearl (1988). Before discussing Pearl's interpretation, some considerations are needed. First, notice that inclusion between propositions over a powerset Boolean algebra coincides with material implication in classical logic. Specifically, the following result is readily provable by properties of Boolean algebra:

Lemma 4.4. *Given a finite non-empty set of possible worlds W, consider its naturally associated Boolean algebra $\langle \wp(W), \cap, \cup, ^-, W, \emptyset \rangle$. Then for any two proposition $X, Y \in \wp(W)$, we have*

$$X \subseteq Y \Leftrightarrow X^- \cup Y = W$$

That is to say, given a possible worlds model $\mathcal{M}$, a proposition X is contained in a proposition Y if the corresponding material implication $X \supset Y$ is valid in $\mathcal{M}$. Now, consider the definition of the Belief function in terms of mass distribution. It can be rewritten as follows:

Remark 4.5 (Pearl 1988)**.** *Consider a non-empty set of possible worlds W and its associated Boolean algebra $\langle \wp(W), \cap, \cup, ^-, \emptyset, W \rangle$. Consider a mass distribution $m : (\wp(W)) \rightarrow [0,1]$ and the induced belief function $Bel : \wp(W) \rightarrow [0,1]$. Then Bel can be reformulated as follows: for all propositions $X \in \wp(W)$,*

$$Bel(X) = \sum_{Y \in \wp(W): Y \subseteq X} m(Y) = \sum_{Y \in \wp(W): X^- \cup Y = W} m(Y)$$

Hence, Pearls' interpretation of a belief function is based upon the idea that $Bel(X)$ represents the degree of *provability* of X. This can be understood by considering that $Bel(X)$ is obtained my summing the mass of all the proposition contained in X, which effectively means summing the masses of all propositions that materially entail X. In other words, $Bel(X)$ quantifies how our available evidence contributes to *proving*, or *materially entailing* the proposition X. We will come back to this point later when discussing the probability of a counterfactual. For now, let us proceed with the study of some technical properties of Belief functions, which will allow us to characterize the probability of a counterfactual.

Example of a Belief function over a Boolean Algebra

Consider the three element Boolean algebra **A** depicted in the following figure.

- A probability distribution, P, assigns weights to the *atoms* of **A**, i.e. α_1, α_2, and α_3.
- A mass distribution m assigns mass to the *elements* of **A**

Consider the probability distribution P and the mass distribution m depicted in the figure below. We have that:

- $P(\alpha_1) = P(\alpha_2) = P(\alpha_3) = \frac{1}{3}$
- $m(\alpha_1) = m(\alpha_2) = m(\neg\alpha_3) = \frac{1}{3}$

By Definition 4.4, we have that $Bel(\neg\alpha_3) = m(\neg\alpha_3) + m(\alpha_2) + m(\alpha_1) = 1$. Clearly Bel is superadditive since $Bel(\alpha_1 \vee \alpha_2) = Bel(\neg\alpha_3) = 1$ but $Bel(\alpha_1) = \frac{1}{3}$ and $Bel(\alpha_2) = \frac{1}{3}$.

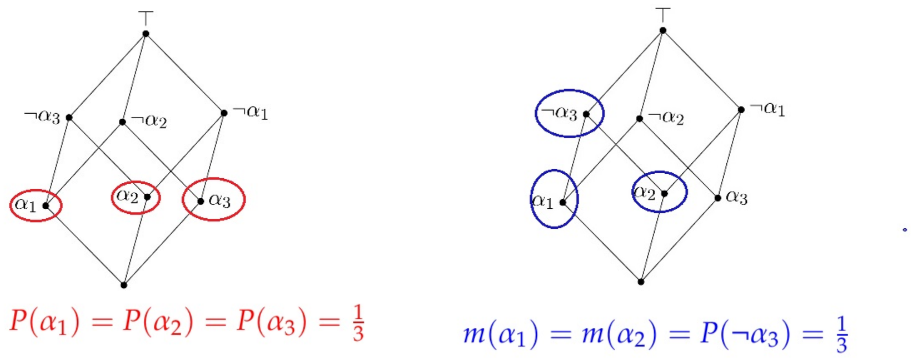


Figure 4.2: The figure depicts a Boolean algebra equipped with a probability distribution on the left, and a Boolean algebra equipped with a mass distribution on the right.

Belief functions are profoundly connected to modal logic. In particular, we can summarize this relationship as follows: classical logic corresponds to probability, while modal logic corresponds to belief functions. The connections between modal logics and belief function have been analyzed from various perspectives. For the purpose of the present thesis, we will highlight some findings from (Hájek 1996; Harmanec, Klir, and Resconi 1994; Resconi, Klir, and Clair 1992) that provide a characterization of belief functions in terms of modal logic. To simplify the notation, we will focus on power-set modal algebras and Kripke frames. However, it is essential to note that the following results hold generally for all modal algebras and their dual frames.

Theorem 4.2 (Hájek 1996; Harmanec, Klir, and Resconi 1994; Harmanec, Klir, and Wang 1996; Resconi, Klir, and Clair 1992)**.** *Consider a Kripke frame $\langle W, R\rangle$ and its associated dual power-set algebra $\langle \wp(W), \cap, \cup, ^{-}, \Box, W, \emptyset\rangle$, by Jónsson-Tarski Duality. Clearly, $\langle \wp(W), \cap, \cup, ^{-}, \Box, W, \emptyset\rangle$ is the Boolean algebra of propositions over W equipped with a modal operator $\Box$ defined as follows: for all $X \in \wp(W)$,*

$$\Box X = \{w \in W \mid R[w] \subseteq X\}$$

with $R[w]$ *being the set of accessible worlds from* w*, is an element of* $\wp(W)$.

Consider any probability distribution $P : W \to [0,1]$. *We can define a mass distribution* $m_P : \wp(W) \to [0,1]$ *as follows: for all* $X \subseteq W$,

$$m_P(X) = \sum_{R[w]=X} P(w)$$

Then m_P *is indeed a a mass distribution and it induces a belief function* $Bel_P : \wp(W) \to [0,1]$ *such that for all* $X \subseteq W$,

$$Bel_P(X) = \sum_{Y \subseteq X} m(Y) = P(\Box X)$$

Proof. The following equalities hold:

$$P(\Box X) = \quad (4.1)$$

$$= \sum_{w \in \Box X} P(w) = \quad (4.2)$$

$$= \sum_{R[w] \subseteq X} P(w) = \quad (4.3)$$

$$= \sum_{R[w]=Y} \sum_{Y \subseteq X} P(w) = \quad (4.4)$$

$$= \sum_{Y \subseteq X} \sum_{R[w]=Y} P(w) = \quad (4.5)$$

$$= \sum_{Y \subseteq X} m_P(Y) = \quad (4.6)$$

$$= Bel_P(X) \quad (4.7)$$

where (4.2) follows from definition of a probability function over an algebra, (4.3) by Jónnsson-Tarski duality, (4.4) and (4.5) follow from the properties of summation, (4.6) from definition of m_P and (4.7) from definition of Bel_P □

The idea behind the above theorem is that if we start with a probability distribution P over a set of possible worlds equipped with an accessibility relation, $\langle W, R \rangle$, i.e. a Kripke frame, we can induce a mass distribution m_P over the algebra of propositions $\wp(W)$ by transferring the original weight of each possible world w, i.e. $P(w)$ to the *set* of its accessible worlds, i.e. $R[w]$. The resulting mass distribution m_P induces a Belief function Bel_P over the algebra of propositions $\wp(W)$, such that the the belief function of each proposition $X \subseteq W$, i.e. $Bel_P(X)$ coincides with the probability of that *necessitated* proposition: $Bel_P(X) = P(\Box X)$. An example will clarify the procedure.

Example: Induced Belief Functions by a Kripke frame

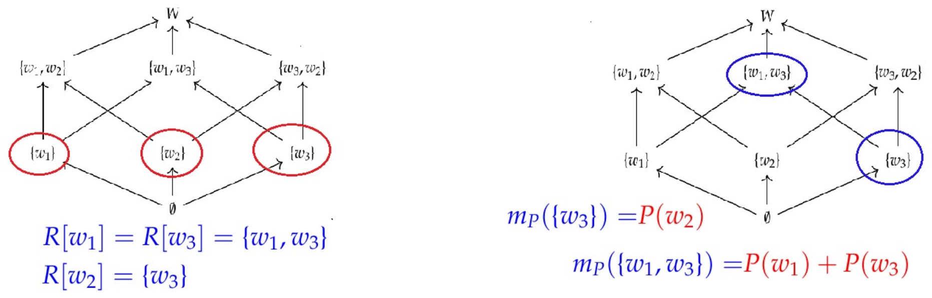


Figure 4.3: The figure depicts an example of how to pass from a probability distribution on a Kripke frame to a corresponding mass distribution

It is interesting to observe that conditions on the accessibility relation R, or, dually, conditions on the $\Box$ operator in the dual modal algebra of $\langle W, R\rangle$, force a certain behavior of the induced belief function. For exemplification, we will show that the seriality constraint over the accessibility relation forces the belief function of contradictory formulas to be 0.

Remark 4.6. *Consider a Kripke model, $\langle W, R, \models\rangle$, equipped with a probability distribution P over W and consider m_P as defined in Theorem 4.2. Assume that R is serial, i.e. for all $w \in W$, $R[w] \neq \emptyset$. Then $m_P(\emptyset) = 0$ and so $Bel_P(\emptyset) = 0$. This implies that impossible formulas, i.e. formulas that cannot be never true at a possible world like classical contradiction, are assigned a belief 0. -*

Proof. Straightforward by the fact that $R[w] \neq \emptyset$ for all w, and so no mass is ever assigned to $\emptyset$. Moreover, contradictory formulas like $\varphi \wedge \neg\varphi$ are such that the proposition they express is equal to the empty set. Hence, their mass and their belief is 0. $\Box$

Now, we have all the ingredients to proceed to a characterization of the probability of a counterfactual.

4.2.2 What is the Probability of a Counterfactual

Recall that in Chapter 2 we have defined Lewis counterfactuals in the language $\mathcal{L}_{LBC_\Box}$ (with finite variables), in the context of $LBC_\Box$-valuations as $\varphi \Box\!\!\rightarrow \psi := \Box(\psi \mid \varphi)$. The following result is readily provable:

Lemma 4.5. *Given a $LBC_{\Box}$ valuation $\langle Perm(Val_{CPL}), R, \vDash\rangle$, consider its dual algebra of propositions $\langle \wp(Perm(Val_{CPL}), \cap, \cup, ^{-}, Perm(Val_{CPL}), \Box, \emptyset, \rangle$ where $\Box$ is defined as in Theorem 4.2. We have that the following hold:*

1. *for all $\Phi \in For_{\mathcal{L}_{LBC_{\Box}}}$, the proposition expressed by Φ, i.e. $[\Phi] = \{\omega \in Perm(Val_{CPL}) \mid \omega \vDash \Phi\} \in \wp(Perm(Val_{CPL}))$;*

2. *for all $\varphi, \psi \in For_{\mathcal{L}}$,*
$$\Box[(\psi \mid \varphi)] = [\Box(\psi \mid \varphi)]$$
where $\Box[(\psi \mid \varphi)]$ is the element obtained by applying $\Box$ to $[(\psi \mid \varphi)]$ in the algebra of propositions $\langle \wp(Perm(Val_{CPL})), \cap, \cup, ^{-}, Perm(Val_{CPL}), \Box, \emptyset, \rangle$

Proof.

1. Straightforward;

2. By semantic conditions of $\Box$, it is easy to show that the following equalities hold:

$$\begin{aligned}
\Box[(\psi \mid \varphi)] &= && (4.8)\\
&= \{\omega \in Perm_{Val_{CPL}} \mid R[\omega] \subseteq [(\psi \mid \varphi)]\} = && (4.9)\\
&= \{\omega \in Perm(Val_{CPL}) \mid \forall \omega' \in R[\omega], \omega' \in [(\psi \mid \varphi)]\} && (4.10)\\
&= \{\omega \in Perm(Val_{CPL}) \mid \forall \omega' \in R[\omega], \omega' \vDash (\psi \mid \varphi\} && (4.11)\\
&= [\Box(\psi \mid \varphi)] && (4.12)
\end{aligned}$$

□

Now, we have all the ingredients to provide a characterization of the probability of a Lewis counterfactual in the logic $\mathbf{VC}^{+}$:

Theorem 4.3. *Consider a $LBC_{\Box}$-valuation $\langle Perm(Val_{CPL}), R, \vDash\rangle$; for any probability distribution P on $Perm(Val_{CPL})$, for all $\varphi, \psi \in For_{\mathcal{L}_{\mid>}^{\Box}}$, we have that:*

$$P(\Box(\psi \mid \varphi)) = Bel_P(\psi \mid \varphi)$$

where Bel_P is defined as in Theorem 4.2.

Proof. The above theorem directly follows from combining the findings from Lemma 4.6 and Theorem 4.2. □

In Chapter 2, we have argued that formulas of the form $\Box(\psi \mid \varphi)$ can be interpreted as Lewis counterfactuals in the contest of $LBC_{\Box}$-valuations, where $(\cdot \mid \cdot)$ is a conditional from $\mathcal{L}_{LBC}$. Then, under this interpretation, the above theorem establishes that *the probability of a Lewis counterfactual is the belief function of the corresponding probabilistic conditional (or Adams conditional)*.

An analogous result can be easily transferred to the case of spherical Kripke models. Recall from Chapter 3 that we have defined Lewis counterfactuals in the language $\mathcal{L}^{\Box}_{>}$ (with possibly infinite variables) in the context of spherical Kripke models satisfying **Centering** from Table 3.1. The following result is readily provable:

Lemma 4.6. *Given a spherical Kripke model $\mathcal{M} = \langle W, \mathcal{S}, R, f, \vDash \rangle$ satisfying* **Centering**, *consider its dual algebra of propositions $\langle \wp(W), \cup, \cap, ^{-}, \Box, \emptyset, W, \rangle$ where $\Box$ is defined as in Theorem 4.2. The following hold:*

1. *for all $\varphi \in For_{\mathcal{L}^{\Box}_{>}}$, the proposition expressed by φ, i.e. $[\varphi] = \{w \in W \mid w \vDash \varphi\} \in \wp(W)$; in particular, $[\Box(\varphi > \psi)] \in \wp(W)$*
2. *for all $\varphi, \psi \in For_{\mathcal{L}^{\Box}_{>}}$, $\Box[\varphi > \psi] = [\Box(\varphi > \psi)]$ where $\Box[\varphi > \psi]$ is the element obtained by applying $\Box$ to $[\varphi > \psi]$ in the algebra $\langle \wp(W), \cap, \cup, ^{-}, W, \Box, \emptyset, \rangle$*

Proof.

1. Straightforward;
2. By semantic conditions of $\Box$, it is easy to show that the following equalities hold:

$$\Box[\varphi > \psi] = \tag{4.13}$$
$$= \{w \in W \mid R[w] \subseteq [\varphi > \psi]\} = \tag{4.14}$$
$$= \{w \in W \mid \text{for all } v \in R[w], v \in [\varphi > \psi]\} = \tag{4.15}$$
$$= \{w \in W \mid \text{for all } v \in R[w], v \vDash \varphi > \psi\} = \tag{4.16}$$
$$= [\Box(\varphi > \psi)] \tag{4.17}$$

□

Now, we have all the ingredients to characterize the probability of a Lewis counterfactual:

Theorem 4.4. *For any spherical Kripke model $\mathcal{M} = \langle W, \mathcal{S}, R, f, \vDash \rangle$ satisfying* **Centering**, *for any probability distribution P on W, for all $\varphi, \psi \in For_{\mathcal{L}^{\Box}_{>}}$, we have that:*

$$P(\Box(\varphi > \psi)) = Bel_P(\varphi > \psi)$$

where Bel_P is defined as in Theorem 4.2.

Proof. The above theorem directly follows from combining the findings from Lemma 4.6 and Theorem 4.2. □

In Chapter 3, we have argued that formulas of the form $\Box(\varphi > \psi)$ can be interpreted as Lewis counterfactuals in the contest of spherical Kripke models satisfying **Centering**, where $>$ is a Stalnaker conditional. Then, under this interpretation, the above theorem establishes that *the probability of a counterfactual is the belief function of the corresponding Stalnaker conditional*.

4.2.3 Imaged Beliefs

We have characterized the probability of Lewis counterfactuals by employing our results from Chapter 2 and Chapter 3, where we explored the definability of Lewis counterfactuals using a modal operator in combination with a Stalnaker (or Adams) conditional However, we can further improve the results from the previous subsection and find a more informative characterization of the probability of a Lewis counterfactual. After all, the hardcore skeptic might still suspect that Theorem 4.4 results from the *trick* of defining a counterfactual in terms of $\Box(\varphi > \psi)$ (or $\Box(\psi \mid \varphi)$). In fact, we have not (yet) provided a characterization the probability of a *genuine* Lewis counterfactual as the primitive connective $\Box\!\!\rightarrow$ within a Lewisian model. n what follows, we will address this skepticism.

Firstly, we present a further equivalent formulation of the belief function from Theorem 4.4. To achieve this, we draw upon some ideas form Dubois and Prade (1994). Recall the generalized imaging procedure from Definition 4.2. Roughly, generalized imaging involves updating an original probability distribution. Suppose we start with a probability distribution P representing our actual belief state. Now, when we learn new evidence φ, we seek to update our probability P in light of this newly acquired information. One way to do so is through generalized imaging. Once more, the idea behind generalized imaging is the following: upon acquiring the information expressed by φ, we require that each $\neg\varphi$-world w loses its weight and redistributes it among its closest φ-worlds, i.e. the worlds in $f(\varphi, w)$. Generalized imaging dictates that a certain fraction of $P(w)$ is transferred to each world in $f(\varphi, w)$. However, the criteria for this redistribution remain arbitrary. Should we transfer the same fraction of $P(w)$ to each world in $f(\varphi, w)$? Or should we transfer to a world v in $f(\varphi, w)$ a fraction of $P(w)$ in proportion to its prior weight? These questions remain unanswered within generalized imaging.

Nevertheless, another approach is possible. Instead of redistributing $P(w)$ among the worlds in $f(\varphi, w)$ we could transfer the weight $P(w)$ to the *whole set* $f(\varphi, w)$, thereby inducing a mass distribution. This idea originates from Dubois and Prade (1994) who claim (the following is adapted to our terminology):

> [...] Instead of sharing $[P(w)$ among the worlds in $f(\varphi, w)]$, a less committed update is to allocate $[P(w)]$ to $f(\varphi, w)$ itself (and none of its subsets). In that case, the imaging process produces a basic probability assignment in the sense of Dempster's view of belief functions. (Dubois and Prade 1994, p.67)

The following results is indeed readily provable:

Lemma 4.7. *Consider any spherical Kripke model $\mathcal{M} = \langle W, \mathcal{S}, R, f, \vDash \rangle$ and probability distribution P on W. Consider its dual algebra of propositions $\langle \wp(W), \cap, \cup, ^{-}, \Box, W, \emptyset \rangle$. Clearly, for all $\varphi \in For_{\mathcal{L}^{\Box}_{>}}$, for all $w \in W$, $f(\varphi, w) \in \wp(W)$.*

For $\varphi \in For_{\mathcal{L}^{\Box}_{\restriction >}}$, consider the function $m^{\varphi}_{P} : \wp(W) \to [0, 1]$ such that for all $X \in \wp(W)$,

$$m^{\varphi}_{P}(X) = \sum_{f(\varphi, w) = X} P(w)$$

m^{φ}_{P} is indeed a mass distribution over $\wp(W)$ and the function $Bel^{\varphi}_{P} : \wp(X) \to [0, 1]$ defined as: for all $X \in \wp(X)$,

$$Bel^{\varphi}_{P}(X) = \sum_{Y \subseteq X} m^{\varphi}_{P}(Y)$$

is indeed a belief function. We refer to Bel^{φ}_{P} as the imaged Belief on φ.

Proof. Straightforward by definition of m^{φ}_{P}

□

Example of the Imaging Procedure from (Dubois and Prade 1994)

Consider a functional Lewisian model $\mathcal{M} = \langle W, f, \vDash \rangle$ (or alternatively a spherical Kripke model) equipped with a probability distribution over W, where:

- there is $w_1 \in W$ such that $P(w_1) = 0.5$
- $f(\varphi, w_1) = \{w_2, w_3\}$

Now, consider the mass distribution m^{φ}_{P} as in Lemma 4.7. The resulting m^{φ}_{P} is depicted as in the following figure. Compare this figure with the Example 4.1.2.

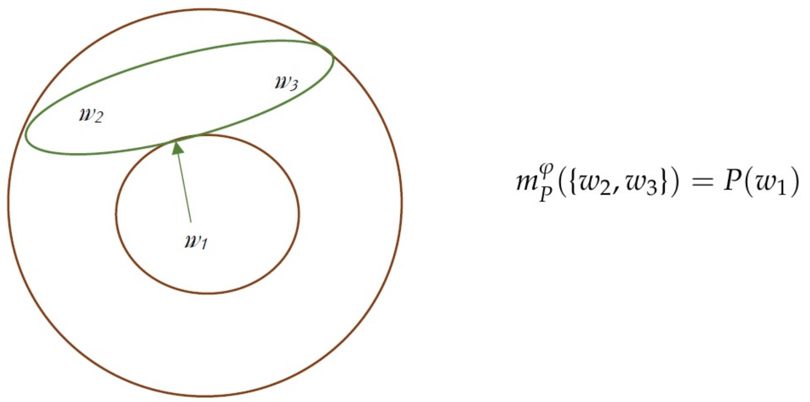


Figure 4.4: A graphical example of the imaging procedure defined by Dubois and Prade (1994).

Intuitively, generalized imaging can be understood as a methods to update our initial probability distribution after acquiring new knowledge. Similarly, imaged belief functions, as proposed by Dubois and Prade (1994), can be regarded as a way to update our initial mass distribution after obtaining new evidence. Let us assume that a probability distribution P models our actual belief state. Now, suppose we learn new evidence encoded by φ. Then each $\neg\varphi$-world w will lose its original weight $P(w)$ and transfer it to the *set* of its most similar φ-worlds $f(\varphi, w)$, resulting in an induced a mass function. We are now ready to prove the following result:

Theorem 4.5. *For any $LBC_{\Box}$-valuation $\langle Perm(Val_{CPL}), R, \models\rangle$, for any probability distribution P over $Perm(Val_{CPL})$, for all $\varphi, \psi \in For_{\mathcal{L}}$, we have that:*

$$P(\Box(\psi \mid \varphi)) = Bel_P(\psi \mid \varphi) = Bel_P^{\varphi}(\psi)$$

where Bel_P^{ψ} is defined as in Lemma 4.7.

Proof. The first equality follows from Theorem 4.3. Then we prove that $P(\Box(\psi \mid \varphi) = Bel_P^{\varphi}(\psi)$. Notice that the following equalities hold by semantic conditions and Lemma 3.4:

$$[\Box(\psi \mid \varphi)] \quad = \tag{4.18}$$

$$= \quad \{\omega \in Perm(Val_{CPL}) \mid R[\omega] \subseteq [(\psi \mid \varphi)]\} = \tag{4.19}$$

$$= \quad \{\omega \in Perm(Val_{CPL}) \mid f(\varphi, \omega) \subseteq [\psi]_{\dashv\vdash_{CPL}}\} \tag{4.20}$$

Hence we have that the following hold:

$$P(\Box(\psi \mid \varphi)) \tag{4.21}$$

$$= \sum_{\omega \models \Box(\psi|\varphi)} P(\omega) = \sum_{f(\varphi,\omega) \subseteq [\psi]_{\dashv\vdash_{CPL}}} P(\omega) \tag{4.22}$$

$$= \sum_{Y \subseteq [\psi]_{\dashv\vdash_{CPL}}} \sum_{f(\varphi,\omega)=Y} P(\omega) \tag{4.23}$$

$$= \sum_{Y \subseteq [\psi]_{\dashv\vdash_{CPL}}} m_P^{\varphi}(Y) = Bel_P^{\varphi}(\psi) \tag{4.24}$$

where (4.22) follows from (4.18)-(4.20); (4.23) follows from properties of summation; (4.24) from definition of m_P^{φ} in Lemma 4.7

□

Once more, all the above results easily transfer to the case of spherical Kripke models in the expanded language $\mathcal{L}_{>}^{\Box}$.

Theorem 4.6. *For any spherical Kripke model $\mathcal{M} = \langle W, \mathcal{S}, R, f, \models \rangle$ satisfying* **Centering***, for any probability distribution P on W, for all $\varphi, \psi \in For_{\mathcal{L}_{|>}^{\Box}}$, we have that:*

$$P(\Box(\varphi > \psi)) = Bel_P(\varphi > \psi) = Bel_P^{\varphi}(\psi)$$

where Bel_P^{ψ} is defined as in Lemma 4.7.

Proof. The first equality follows from Theorem 4.4. Then, we prove that $P(\Box(\varphi > \psi)) = Bel_P^{\varphi}(\psi)$. Notice that the following equalities hold by semantic conditions and Lemma 3.4:

$$[\Box\varphi > \psi] \quad = \tag{4.25}$$

$$= \quad \{w \in W \mid R[w] \subseteq [\varphi > \psi]\} = \tag{4.26}$$

$$= \quad \{w \in W \mid f(\varphi, w) \subseteq [\psi]\} \tag{4.27}$$

Hence we have that the following hold:

$$P(\Box(\varphi > \psi)) \tag{4.28}$$

$$= \sum_{w \models \Box(\varphi > \psi)} P(w) = \sum_{f(\varphi,w) \subseteq [\psi]} P(w) \tag{4.29}$$

$$= \sum_{Y \subseteq [\psi]} \sum_{f(\varphi,w)=Y} P(w) \tag{4.30}$$

$$= \sum_{Y \subseteq [\psi]} m_P^{\varphi}(Y) = Bel_P^{\varphi}(\psi) \tag{4.31}$$

where (4.29) follows from (4.25)-(4.27); (4.30) follows from properties of summation; (4.31) from definition of m_P^{φ} in Lemma 4.7

□

Notice that the above proofs can straightforwardly be adapted to Lewisian models as well:

Corollary 4.2. *For any functional Lewisian model $\mathcal{M} = \langle W, f, \vDash \rangle$ satisfying* **Centering**, *for any probability distribution P on W, for all $\varphi, \psi \in For_{\mathcal{L}_{\Box\!\!\rightarrow}}$, we have that:*

$$P(\varphi \Box\!\!\rightarrow \psi) = Bel_P^{\varphi}(\psi)$$

where Bel_P^{ψ} is defined as in Lemma 4.7.

Example: Belief function induced by a functional Lewisian model

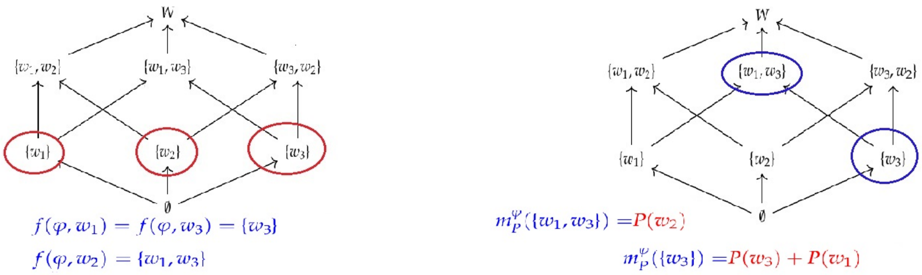


Figure 4.5: The figure depicts an example of how to pass from a probability distribution on the algebra induced by a spherical Kripke frame (or, equivalently, by a functional Lewisian model) to a corresponding mass distribution induced by the selection function f.

Analogously to Remark 4.6, it is possible to show that axioms of **VC**-logic, or, equivalently, constraints over the selection function f in Table 2, force some properties of the corresponding imaged belief functions. Here is an example:

Remark 4.7. *Consider a functional Lewisian model $\langle W, f, \vDash\rangle$ satisfying* **Centering** *equipped with a probability distribution P over W. Then, for every $\varphi \in For_{\mathcal{L}_{\Box\rightarrow}}$, the following holds*

$$Bel_P^{\varphi}(\varphi) = 1$$

Namely the imaged belief on φ forces that φ is certain.

Proof. Straightforward by the fact that in every functional Lewisian model $\langle W, f, \vDash\rangle$, for all $\varphi \in For_{\mathcal{L}_{\Box\rightarrow}}$, for all $w \in W$, we have that $f(\varphi, w) \subseteq [\varphi]$. By Lemma 4.7, this implies that for all $X \in \wp(W)$, if $m_P^{\varphi}(X) > 0$, then $X \subseteq [\varphi]$. Hence $\sum_{X\subseteq[\varphi]} m_P^{\varphi}(X) = 1$ □

Observe that the above results hold for *all* Lewisian models and so for all kinds of variably strict conditionals. Therefore, we have provided a faithful characterization of the proposition expressed by a Lewis counterfactual that can also be extended to all Lewis variably strict conditionals, thereby providing an answer to question (P1) in the introduction.

4.3 A Probabilistic Look at Lewis Counterfactuals

In this section, we present some philosophical considerations regarding the results we have proven above. Specifically, we show how our characterization of the probability of counterfactuals has a natural interpretation that shed new light on the understanding of Lewis counterfactuals as well. We will apply Pearl's interpretation of belief functions in terms of the notion of *provability* to the case of spherical Kripke models. However, it's important to note that all the following results can be easily adapted to the case of Lewis frames and Lewis algebras.

Recalling Remark 4.5, we observe that it can be readily adapted to the case of spherical Kripke models:

Remark 4.8. *Given a spherical Kripke model $\mathcal{M} = \langle W, \mathcal{S}, R, f, \vDash\rangle$ satisfying* **Centering**, *consider its naturally associated algebra of propositions $\langle \wp(W), \cap, \cup, ^-, W, \Box, \emptyset, \rangle$, a a probability distribution P on W, the induced mass distribution m_P over W and the corresponding belief function Bel_P. We have that for all $\Box(\varphi > \psi) \in For_{\mathcal{L}^{\Box}_{>}}$,*

$$P(\Box(\varphi > \psi)) = Bel_P(\varphi > \psi) = \sum_{Y\subseteq[\varphi>\psi]} m_P(Y) = \sum_{Y^-\cup[\varphi>\psi]=W} m_P(Y)$$

So, it is plausible to interpret material implication between two proposition as a degree of *provability* according to Pearl (1988): $X \supset Y = X^- \cup Y = W$ if and only if $X \subseteq Y$, meaning that X is contained in Y when X materially entails Y. Recall that for all proposition X, $m_P(X)$ can be interpreted as the *strength* of the evidence we have in favor of proposition X, or, more simply, $m_P(X)$ quantifies the available evidence supporting X. Hence, under Pearl's interpretation, $P(\Box(\varphi > \psi)) = Bel_P(\varphi > \psi)$ can be interpreted as quantifying *how much* $\varphi > \psi$ is implied by our available evidence. Roughly, it quantifies how much $\varphi > \psi$ is provable from our available evidence, or, under a more frequentist-like interpretation, how *frequently* $\varphi > \psi$ is provable from our available evidence. In comparison to the intuitive interpretation of probability, where $P(\varphi > \psi)$ quantifies the chances that $\varphi > \psi$ is *true*, $P(\Box(\varphi > \psi)) = Bel_P(\varphi > \psi)$ is quantifying the chances that $\varphi > \psi$ is *provable*. An analogous reasoning can be applied to the case of imaged belief. Recall that $Bel_P(\varphi > \psi) = Bel_P^{\varphi}(\psi)$. Hence, quantifying the degree of provability of $(\varphi > \psi)$, i.e. $Bel_P(\varphi > \psi)$, amounts to quantifying the degree of provability of ψ under the *imaged* assumption that φ holds, i.e. $Bel_P^{\varphi}(\psi)$.

These considerations seem to support the view mentioned in Chapter 2 and Chapter 3 that the $\Box$ operator in our language $\mathcal{L}_{>}^{\Box}$ can be interpreted as a provability modality in the contest of spherical Kripke models. These observations can be easily be applied to the case of primitive Lewis counterfactuals in the language $\mathcal{L}_{\Box\!\!\rightarrow}$. $P(\varphi \Box\!\!\rightarrow \psi) = Bel_P^{\varphi}(\psi)$ can be interpreted as the degree of provability of ψ after imaging φ.

One last comment is due concerning imaged belief functions. First, it would be illuminating to view our results through this schema:

$$\begin{array}{ccc} P(\varphi > \psi) & \longrightarrow & P^{\varphi}(\psi) \\ \downarrow & & \downarrow \\ Bel_P(\varphi > \psi) & \longrightarrow & Bel_P^{\varphi}(\psi) \end{array}$$

Figure 4.6: A picture summarizing our results concerning the probability of Lewis counterfactuals and their connections with the probability of Stalnaker conditionals.

Lewis (1976) provided a characterization of the probability of a Stalnaker conditional in terms of imaged probability, showing that *the probability of a Stalnaker conditional coincides with the corresponding imaged probability*. Our findings can be seen a generalization of Lewis' results to the case of belief functions, where *the belief of a Stalnaker conditional coincides with the corresponding imaged belief function*. This observation leads to another possible answer to question (L1a) regarding the relationship between Stalnaker conditionals and Lewis counterfactuals from a probabilistic perspective. While Stalnaker was interested in accounting for the truth-conditions of the conditional > in his theory, Lewis' theory accounts for the *provability-conditions* of the same conditional. Thus, in a sense, the above observations, along with those from Chapter 2 and Chapter 3, seem to suggest that Stalnaker's and Lewis' theories are about the same conditional dependence expressed by >, but approach it from two different perspectives. Moreover, the probability of Stalnaker conditionals and Lewis counterfactuals mirror each other, as the probability of a Stalnaker conditional coincides with the corresponding imaged probability, while the probability of a Lewis counterfactual coincide with the corresponding imaged belief function.

This probabilistic account of Lewis counterfactuals also offers a new interpretation of the meaning of the axioms of Lewis' logic **VC** (or **KVC** in the case of spherical Kripke models). For example, consider the axiom $\varphi \mathrel{\Box\!\!\rightarrow} \varphi$. This axiom implies that the probability of counterfactuals with the same formula as both antecedent and consequent is always equal to 1, i.e. $P(\varphi \mathrel{\Box\!\!\rightarrow} \varphi) = 1$ (see also Remark 4.7). Correspondingly, the imaged belief of φ on φ, $Bel_P^{\varphi}(\varphi)$ is always 1, i.e. $Bel_P^{\varphi}(\varphi) = 1$. Analogously, also other axioms can be interpreted along the same line: they provide constraints over the imaging procedure applied to belief function. For instance, consider the complex axiom $((\varphi \vee \psi) \mathrel{\Box\!\!\rightarrow} \varphi) \vee ((\varphi \vee \psi) \mathrel{\Box\!\!\rightarrow} \psi) \vee (((\varphi \vee \psi) \mathrel{\Box\!\!\rightarrow} \delta) \leftrightarrow ((\varphi \mathrel{\Box\!\!\rightarrow} \delta) \wedge (\psi \mathrel{\Box\!\!\rightarrow} \delta)))$. This axiom characterizes the following constraint over the selection function:

$$f(\varphi \vee \psi, w) \subseteq [\varphi] \text{ or } f(\varphi \vee \psi, w) \subseteq [\psi] \text{ or } f(\varphi \vee \psi, w) = f(\varphi, w) \cup f(\psi, w)$$

The selection function, according to Lemma 4.7, induces the corresponding mass function that generates the imaged belief. Hence, the above constraint on f establishes how the imaged mass $m_P^{\varphi \vee \psi}$ must be allocated. Specifically, it states that the weight of $P(w)$, under the imaging assumption of $\varphi \vee \psi$, will be transferred either to a subset of $[\varphi]$ (since $f(\varphi \vee \psi, w) \subseteq [\varphi]$), or to a subset of $[\psi]$ (since $f(\varphi \vee \psi, w) \subseteq [\psi]$), or to the union of $f(\varphi, w)$ and $f(\psi, w)$. This implies that the weight of $P(w)$ will only be transferred to the whole proposition $[\varphi \vee \psi]$ in extreme cases. A similar reasoning applies to other constraints of the selection function. Under this interpretation, Lewis' logic of counterfactuals (or variably strict conditionals) can be regarded as the logic characterizing special types of imaged belief functions.

4.4 Conclusions

The results presented above provide an answer to question (P1) regarding the characterization of the probability of a Lewis counterfactual. Initially, we showed how generalized imaging fails to account for the probability of Lewis counterfactuals. Subsequently, our translation from Chapters 2 and 3, combined with classical results connecting belief functions and modal logic, guided us in characterizing the probability of a counterfactual in terms of a belief function. We demonstrated that the probability of a counterfactual $\Box(\varphi > \psi)$ ($\Box(\psi \mid \varphi)$) corresponds to the belief function of the corresponding Stalnaker conditional $(\varphi > \psi)$ (*LBC*-conditional $(\psi \mid \varphi)$). Furthermore, we established that this characterization is not tied to our translation of Lewis counterfactuals, but it naturally extends to primitive counterfactuals, $\Box\!\!\rightarrow$, within Lewis language. This result also provides new insights into the interpretation of Lewis counterfactuals, suggesting that the probability of a Lewis counterfactual quantifies the degree of *provability* of the corresponding Stalnaker conditional. This aligns with our previous considerations from Chapter 2 and Chapter 3 concerning the interpretation of the modality expressed by $\Box$ as a *provability* modality.

The belief function corresponding to the probability of Lewis counterfactuals can, in turn, be characterized using a special non-Bayesian updating procedure suggested by Dubois and Prade (1994), resulting in an *imaged* belief function. This type of updating method closely mirrors generalized imaging from classical probability, but it has not been extensively explored. Lewis' semantics for counterfactuals seems closely related to this kind of imaged belief functions, in the sense that characteristics axioms of variably strict conditional logics or, dually, properties of the selection function, constraint the corresponding imaged belief functions. This seems to suggest that Lewis' variably strict conditional logics can serve as the logical counterpart of an imaging-like updating method for belief functions.

Chapter 5

Probability of Counterfactuals within Causal Modeling Semantics

In all the previous chapters we have been working within the framework of possible worlds semantics and algebraic semantics for counterfactuals. However, as we mentioned in the introduction, there are also many alternative semantic accounts for counterfactuals. For instance, the dominant paradigm in computer science and formal epistemology is *causal modeling semantics* (CMS). In this section, we address question (P2) posited in the introduction and propose a new method for computing the probability of counterfactuals with complex antecedents within the causal modeling framework.

5.1 Background

In this section we introduced causal modeling semantics for counterfactuals, mention some of its shortcoming, and review Briggs' (2012) proposal to expand standard causal modeling semantics.

The basic idea behind CMS is that a counterfactual $A \mathrel{\Box\!\!\to} B$ is interpreted relative to a causal model $\mathcal{M}$[1]. It is true if an *intervention* forcing the event A in $\mathcal{M}$ also yields B, and false if this intervention does not yield B (Pearl 2000, 2017). This proposal, which relies on causal models as a graphical tool for reasoning and inference, is elaborated in Galles and Pearl (1998). On this account, the "probability of counterfactual statements" (Pearl 2000, p. 205) is

[1]Observe that we are using the same symbol from $\mathcal{L}_{\Box\!\!\to}$ to denote the counterfactual connective, however it is important to keep in mind the the logic underlying causal modeling semantics is different from Lewis' logic **VC**, as well as the truth conditions of a counterfactual within a causal model differ from the standard Lewisian truth conditions.

interpreted as the probability that after an intervention on A (written $do(A)$), B will hold: $P(A \mathbin{\Box\!\!\rightarrow} B) = P_{do(A)}(B)$.

The divergences and convergences of CMS and Lewis' semantics for counterfactuals have been studied from various angles. Pearl (2000, pp. 72-73) shows that a particular type of imaging is equivalent to an intervention on A that is represented by the *do*-operator. It is agreed, however, that standard CMS and Lewis account are different in at least one crucial respect: they assign truth conditions to different classes of counterfactuals. The Lewisian framework assigns truth values—and probabilities—to counterfactuals $A \mathbin{\Box\!\!\rightarrow} B$ with *arbitrary* antecedents, regardless of their logical complexity, since for any sentence A, the set of closest possible A-worlds is well-defined.

By contrast, Standard CMS, as developed in Galles and Pearl (1998), cannot account for the truth conditions or probability of counterfactuals with disjunctive antecedents of the form $(A \vee B) \mathbin{\Box\!\!\rightarrow} C$, e.g., "if it had rained *or* there had been riots, the football match would have been cancelled". The reason is that it is simply not clear which intervention corresponds to the logical *disjunction* of two atomic interventions. In other words, while CMS has a strong theoretical motivation and a history of successful applications, it has limited expressive power.

In this chapter we aim to close the above gap: building on Briggs' 2012 pioneering work on expanding CMS and ideas from truthmaker semantics (Fine 2016, 2017), we propose a CMS-based account for evaluating the probability of counterfactuals with disjunctive antecedents. Specifically, we propose to evaluate the probability of $(A \vee B) \mathbin{\Box\!\!\rightarrow} C$ as the weighted probability of C in all submodels that *truthmake* $A \vee B$. The relative weights of the submodels are determined by their distance to the original model, based on a metric developed in Eva, Stern, and Hartmann (2019). This procedure extends CMS to calculating the probability of counterfactuals with arbitrary Boolean compounds of atomic formulas in the antecedent. We also show that the predictions of our account are superior to the ones obtained by Lewis' imaging procedure.

5.1.1 Causal Modeling: Syntax and Semantics

First, we need to introduce causal models, using a running example (simplified from Pearl 2000) that will accompany us throughout our discussion. It involves four Boolean variables, whose values are represented by the numbers zero and one.

Example: Execution Scenario

> A prisoner is condemned to death and led to the execution court. He stands in front of two soldiers, who will fire at the captain's signal. If at least one of the soldiers fires, the prisoner dies. The captain gives the signal ($C = 1$), the two soldiers fire ($X = 1$, $Y = 1$), and the prisoner dies ($D = 1$).

The main ingredients of this causal model are:

- a set of variables $\mathcal{V} = \{C, X, Y, D\}$;
- the set of structural equations that describe their causal dependencies: $\mathcal{S} = \{X = C, Y = C, D = max(X, Y)\}$

This means that the executioners fire if the captain gives the signal and the prisoner dies if one of the two executioners fires. The dependencies can also represented graphically, as in Figure 5.1 below.

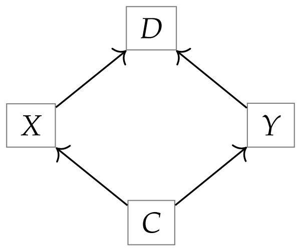


Figure 5.1: Causal graph for the prisoner execution story. C stands for the captain (not) firing, X, Y for the soldiers (not) shooting, D for the prisoner dying/living.

- The *parents* $PA(V)$ of a variable V are simply the variables from which there is an arrow into V.
- For example, C is the only parent of X and Y, and X and Y are the parents of D.
- Structural equations describe the value of a variable as a function of the value of its parents.

Before defining a causal model, we need to specify our languages. We mainly work through two different language levels:

Definition 5.1. *An atomic formula in our language has the form $V = v$, meaning that the variable V takes a certain value v. "At" denotes the sets of our atomic formulas. Starting from At, we can define two different languages. Uppercase Latin letters $A, B, C, \ldots$ are used to indicate formulas in our languages:*

1. *a language $\mathcal{L}_{CMS}$ containing connectives $\wedge, \vee$, and $\neg$ consisting a simple classical language over the set of our atomic formulas At.*
    - *formulas in $\mathcal{L}_{CMS}$ are defined as usual: every $V = v \in At$ is a formula of $\mathcal{L}_{CMS}$, if A, B are formulas of $\mathcal{L}_{CMS}$, then also $\neg A, A \wedge B, A \vee B$ are, and noting else is a formula of $\mathcal{L}_{CMS}$*
    - *$For_{\mathcal{L}_{CMS}}$ indicates the set of formulas of $\mathcal{L}_{CMS}$.*
2. *a more complex language,referred to $\mathcal{L}^{\Box\!\!\to}_{CMS}$, which consists in expanding $\mathcal{L}_{CMS}$ with a binary connective $\Box\!\!\to$, read as the counterfactual conditional connective. A formula in $\mathcal{L}^{\Box\!\!\to}_{CMS}$ is inductively defined as follows:*
    - *if $A \in For_{\mathcal{L}_{CMS}}$, then A is a formula of $\mathcal{L}^{\Box\!\!\to}_{CMS}$*
    - *if A, B are formulas of $\mathcal{L}_{CMS}$, then $A \Box\!\!\to B$ is a formula of $\mathcal{L}^{\Box\!\!\to}_{CMS}$*
    - *if A, B are formulas of $\mathcal{L}^{\Box\!\!\to}_{CMS}$, then so are $\neg A, A \wedge B, A \vee B$*
    - *nothing else is a formula of $\mathcal{L}_{CMS}$*

    Moreover, $For_{\mathcal{L}^{\Box\!\!\to}_{CMS}}$ denotes the set of formulas in $\mathcal{L}^{\Box\!\!\to}_{CMS}$. Basically, nested occurrences of $\Box\!\!\to$ are not allowed in $\mathcal{L}^{\Box\!\!\to}_{CMS}$.

In general, a causal model can be defined as follows:

Definition 5.2. *A causal model is a triple $\mathcal{M} = \langle \mathcal{V}, \mathfrak{S}, a \rangle$ where:*

- *$\mathcal{V}$ is a non-empty finite set of variables $\mathcal{V} = \{V_1, V_2, ..., V_n\}$;*
- *$\mathfrak{S}$ is a set of structural equations, where each element has the form $V = f_V(V_{i_1}, V_{i_2}, \ldots, V_{i_n})$ and $PA(V) = \{V_{i_1}, \ldots, V_{i_n}\}$ (i.e., each structural equation defines the value of V uniquely by the value of its parents; no cycles are allowed);*
- *$a : \mathcal{V} \to \mathcal{R}(\mathcal{V})$ is a function assigning an actual value to each variable V, in a way that is consistent with the range of V and the structural equations.*

The last part, the assignment of actual values, is not necessarily required for making predictions with causal models, but it is crucial when we want to use them for counterfactual reasoning. Some additional terminology will be useful: when a variable V_1 is connected to another variable V_2 via a sequence of directed arrows from V_1 into V_2, we say that V_2 is a *descendant* of V_1. For instance, in Figure 5.1, D is a descendant of C, X and Y. As in Briggs (2012), we will restrict our attention to models not containing any loops, i.e., models where there is no sequence of arrows connecting a variable to itself. Moreover, in a causal model, we say that a variable is *exogenous* when it has no parents (e.g., C in Figure 5.1) and *endogenous* when it is not exogenous, so that its value can be determined by the value of other variables in the model (e.g., X, Y and D in Figure 5.1).

Now, we introduce the notion of an *intervention* on a causal model.

Definition 5.3. *Given a causal model $\mathcal{M} = \langle \mathcal{V}, \mathfrak{S}, a \rangle$, an intervention $do(V = v)$ on a causal model breaks the dependency of V on its parents via the structural equations (i.e., it eliminates all arrows into V) and assigns the value $V = v$ to it.*

The intervention generates a causal sub-model $\mathcal{M}'$ where the formula $V = v$ is true and the structural equation f_V is no longer part of the causal model: the variable V now depends on the intervention, but no longer depends on its parents.

The above idea can be generalized to conjunctions of interventions: the intervention $do(V_1 = v_1, V_2 = v_2, \ldots, V_n = v_n)$ generates a sub-model $\mathcal{M}' = \langle \mathcal{V}', \mathfrak{S}', a' \rangle$ of $\mathcal{M}$ such that:

- *$\mathcal{V}' = \mathcal{V}$, i. e. $\mathcal{M}'$ has the same variables as $\mathcal{M}$;*
- *$\mathfrak{S}' = \mathfrak{S} \setminus \{f_{V_1}, \ldots, f_{V_n}\}$;*
- *$a' : \mathcal{V} \setminus \{V_1, V_2, \ldots, V_n\} \to \mathcal{R}(\mathcal{V})$ assigns actual values to the variables not affected by the intervention, in line with the structural equations in $\mathfrak{S}'$.*

On an intuitive level, an intervention on a causal model manipulates some variables, forces them to take a certain value and breaks the causal mechanism between them and their parents. For an example, consider the causal model of the execution story depicted above; we want to know what would have happened if the two executioners had not fired $(X = 0 \wedge Y = 0)$. The answer is given by the intervention $do(X = 0, Y = 0)$ which would generate the model in Figure 5.2.

Our intervention has broken the causal mechanism that links C to X and Y, and we have forced X and Y to value zero. What happens to D now? It continues to be determined by the structural equation $D = max(X, Y)$, but $X = 0$ and $Y = 0$ as a result of our intervention, hence $D = max(0, 0) = 0$. And so the prisoner will live.

The intuitive counterfactual reasoning within a causal model seems to run along these lines: in order to know *what would have happened* to the prisoner *had the executioners not fired*, we perform an intervention on the latter and

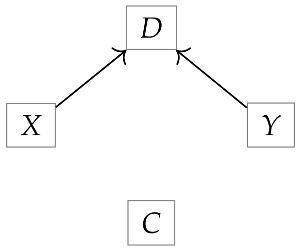


Figure 5.2: Causal graph for the prisoner execution story, where we intervene on X and Y and break the dependency on the captain's signal C.

see how it would have affected the prisoner, according to the known causal mechanisms, and without changing any facts that are causally independent of the executioners' actions. CMS explicates this line of thought in a mathematically precise way (e.g., Pearl 2000, p. 205). A counterfactual supposition in a causal model amounts to an external action on the model that enforces that supposition with a minimal change in the structure of the network (i.e., an intervention).

Specifically, a counterfactual sentence of the form $(A_1 \wedge A_2 \wedge ... \wedge A_n) \Box\!\!\rightarrow B$ is true at a causal model $\mathcal{M}$ that contains $A_1, \ldots, A_n$ and B as variables if and only if at the causal model $\mathcal{M}'$ generated by the intervention $do(A_1 = 1, A_2 = 1, \ldots, A_n = 1)$ on $\mathcal{M}$, we also have $B = 1$.[2] For instance, the counterfactual "if the two executioners hadn't fired, then the prisoner would not have died" is true at the causal model of the execution story since, as we have seen above, $D = 0$ holds in the new submodel after performing the intervention $do(X = 0, Y = 0)$.

Notice that an intervention of the form $do(A)$ is only defined when A is an atomic formula or a conjunction of atomic formulas. This imposes a restriction on the class of counterfactuals that standard CMS can account for: only counterfactuals of the form $(A_1 \wedge A_2 \wedge ... \wedge A_n) \Box\!\!\rightarrow B$ can assume a truth value. CMS does not provide truth conditions for counterfactuals with logically complex antecedents. For instance, we cannot say whether the counterfactual "if one of the two executioners hadn't fired, then the prisoner would not have died" ($(X = 0 \vee Y = 0) \Box\!\!\rightarrow D = 0$) is true or false at the causal model of the execution story. This limitation is due to the fact that the *disjunctive intervention* $do(X = 0 \vee Y = 0)$ is not defined (see also Pearl 2017). Intuitively, there is more than one possible realization of $do(X = 0 \vee Y = 0)$: we could manipulate X, Y, or both variables at the same time (compare Briggs 2012; Günther 2017; Sartorio 2006). Each of the three interventions $do(X = 0)$, $do(Y = 0)$ and $do(X = 0, Y = 0)$ would be a good candidate for an intervention that brings

[2] As before, we use $A_1 = 1$ for expressing that the Boolean variable A_1 takes the value "true".

about the state "$X = 0$ or $Y = 0$". But their effects on $D = max(X, Y)$ differ. For the intervention $do(X = 0)$ and $do(Y = 0)$, the prisoner would still die (since the other soldier fires) but for the intervention $do(X = 0, Y = 0)$, he would live. Thus, if *just one* executioner hadn't fired, the prisoner would have died anyway; if *both* hadn't fired, he would live. So, in the end, standard CMS as presented in Galles and Pearl (1998) and Pearl (2000) does not provide a unique answer to the question of evaluating counterfactuals with disjunctive antecedents. This is arguably a disadvantage of CMS with respect to Lewis' account, where selection functions provide definite answers to the question of which are the relevant worlds for evaluating counterfactuals, and how the results need to be combined (e.g., Lewis demands that the consequent holds in all nearest possible worlds where the antecedent is true). In order to overcome this shortcoming, Briggs (2012) has proposed an extension of CMS that we present in the next section.

5.1.2 Truthmaker Semantics for Causal Modeling

Briggs' extension of CMS relies on truthmaker semantics (TMS), a semantic framework developed in a series of recent publications by Kit Fine (2016, 2017). The idea underlying TMS is that of an *exact truthmaker* of a sentence A, namely something in the world which is *responsible* and *wholly relevant* for the truth of A. One of the motivations behind truthmaker semantics is to be able to draw hyperintesional distinctions between propositions, i.e., to distinguish propositions that would be otherwise identical in the classical possible worlds framework, like p and $p \vee (p \wedge q)$, or tautologies like $p \vee \neg p$ and $q \vee \neg q$. More precisely, the fundamental structure in TMS is that of a state space $\langle S, \sqsubseteq \rangle$ where S is a non-empty set of states which stand for portions of reality (e.g., facts, events, individuals etc.), and $\sqsubseteq$ is a partial order over S that can be understood as parthood relation between the elements in S. We can then define an operation $\sqcup$ of *fusion* between states: given two states s and t, their fusion of $s \sqcup t$ is the least upper bound of the set $\{s, t\}$. We can equip a state space with interpretation functions so as to define a relation of *exact truthmaking* and *exact falsemaking* between sentences and states so that those states can be truthmakers or falsemakers of formulas (for more details on TMS see for instance Fine (2017)). $s \Vdash A$ ($s \dashv\!| A$) indicates that s is an exact truthmaker (falsemaker) of A.

Briggs (2012) shows how truthmaker semantics can expand the scope of CMS. More precisely, causal modeling semantics can be expanded as follows:

Notation 5.1.

- *An intervention $do(A)$ is admissible on a causal model $\mathcal{M}$ when it does not perform two inconsistent value assignments to the same variable, like $do(V_1 = 0 \wedge V_1 = 1)$.*

Definition 5.4. *For a causal model $\mathcal{M} = \langle \mathcal{V}, \mathcal{S}, a \rangle$, we can define the* set of submodels *of $\mathcal{M}$ generated by any intervention $do(A)$ as $S(\mathcal{M}) = \langle S, \sqcup \rangle$ where*

- *S is the set of submodels of $\mathcal{M}$ generated by any admissible intervention $do(A)$;*
- *$\mathcal{M}[A]$ indicates the submodel generated by performing $do(A)$ on $\mathcal{M}$;*
- *$\sqcup$ is an operation of* fusion *among the models in S defined by $\mathcal{M}[A] \sqcup \mathcal{M}[B] := \mathcal{M}[A \wedge B]$.*

In other words, the fusion of the two submodels $\mathcal{M}[A]$ and $\mathcal{M}[B]$, defined by the interventions $do(A)$ and $do(B)$, corresponds to the submodel defined by the fusion of the two interventions, where the fusion of two interventions is simply the intervention that encodes both, i.e. the conjunctive intervention of both of them. We assume that only logically consistent fusions are allowed. For instance, let X be a variable in a model $\mathcal{M}$ which stands for the status of the light: $X = 0$ means that the light is off, and $X = 1$ means that the light is on. It is then impossible to fuse $\mathcal{M}[X = 0]$ and $\mathcal{M}[X = 1]$, because their fusion would yield a model where the light is both on and off, or in other words, the intervention $do(X = 0 \wedge X = 1)$ is not admissible.

Under this expansion of causal model we can define new truth-making conditions for formulas in $For_{\mathcal{L}_{CMS}}$, i.e. not containing the counterfactual conditional connective:

Definition 5.5. *For a model $\mathcal{M}$, consider its space of proper submodels $S(\mathcal{M}) = \langle S, \sqcup \rangle$ where $\mathcal{M} \notin S$. We can inductively define relations of truthmaking $\Vdash \subseteq S \times For_{\mathcal{L}_{CMS}}$ and and falsemaking $\dashv\mid \subseteq S \times For_{\mathcal{L}_{CMS}}$ between any member s of S and formulas in the language as follows:*

$$
\begin{array}{lll}
s \Vdash V = v & \Leftrightarrow & s = \mathcal{M}[V = v] \\
s \dashv\mid V = v & \Leftrightarrow & s = \mathcal{M}[V = v'] \textit{ for some } v \neq v' \\
s \Vdash \neg A & \Leftrightarrow & s \dashv\mid A \\
s \dashv\mid \neg A & \leftrightarrow & s \Vdash A \\
s \Vdash A \wedge B & \Leftrightarrow & \textit{for some } t, u \ (t \Vdash A, u \Vdash B \textit{ and } s = t \sqcup u) \\
s \dashv\mid A \wedge B & \Leftrightarrow & s \dashv\mid A,\ s \dashv\mid B,\ \textit{or } s \dashv\mid A \vee B \\
s \Vdash A \vee B & \Leftrightarrow & s \Vdash A,\ s \Vdash B,\ \textit{or } s \Vdash A \wedge B \\
s \dashv\mid A \vee B & \Leftrightarrow & \textit{for some } t, u \ (t \dashv\mid A, u \dashv\mid B \textit{ and } s = t \sqcup u)
\end{array}
$$

where $s \Vdash A$ means that s truthmakes (=is a truthmaker of) A.

We say that a state s is a truthmaker of $V = v$ if and only if it corresponds to the submodel defined by the intervention $do(V = v)$, and a falsemaker of $V = v$ if and only if it corresponds to the submodel defined by an intervention that sets V to a value different from v. Since states in $S(\mathcal{M})$ can be identified with interventions, we can say, for simplicity, that an intervention $do(V_1 = v_1, ..., V_n = v_n)$ on $\mathcal{M}$ truthmakes a formula A if and only if $\mathcal{M}[V_1 = v_1, ..., V_n = v_n]$ is a truthmaker of A.

Evidently, *s* falsemakes *A* iff *s* is a truthmaker of $\neg A$. State *s* truthmakes a *conjunction* of variable assignments iff it is the fusion of two states that truthmake the two individual assignments—in other words, if and only if *s* is the causal submodel defined by the intervention that assigns the right values to both variables. Finally, *s* is truthmaker of a *disjunction* of variable assignments iff it truthmakes one of the two assignments, or its conjunction. This interpretation of truthmaking a disjunction is also at the center of Briggs' (and our own) proposal for expanding CMS.

We can now give inductively defined truth conditions for formulas of $\mathcal{L}^{\Box\!\!\rightarrow}_{CMS}$, including simple counterfactuals.

Definition 5.6. *A formula in $For_{\mathcal{L}^{\Box\!\!\rightarrow}_{CMS}}$ is true at a causal model $\mathcal{M} = \langle \mathcal{V}, \mathfrak{S}, a \rangle$ according to the following inductive clauses:*

$$
\begin{array}{lll}
\mathcal{M} \vDash V = v & \Leftrightarrow & a(V) = v \\
\mathcal{M} \vDash \neg A & \Leftrightarrow & \mathcal{M} \nvDash A \\
\mathcal{M} \vDash A \wedge B & \Leftrightarrow & \mathcal{M} \vDash A \text{ and } \mathcal{M} \vDash B \\
\mathcal{M} \vDash A \vee B & \Leftrightarrow & \mathcal{M} \vDash A \text{ or } \mathcal{M} \vDash B \\
\mathcal{M} \vDash A \mathrel{\Box\!\!\rightarrow} B & \Leftrightarrow & \text{for every } s \text{ in } S(\mathcal{M}) \text{ such that } s \Vdash A, s \vDash B
\end{array}
$$

Thus, a counterfactual $A \mathrel{\Box\!\!\rightarrow} B$ is true at a causal model $\mathcal{M}$ if and only if *B* is true at all the members of $S(\mathcal{M})$ that truthmake *A*. Consider again the execution example and the counterfactual "if one of the two executioners had not fired, then the prisoner would not have died". We can formalize this counterfactual as $(X = 0 \vee Y = 0) \mathrel{\Box\!\!\rightarrow} D = 0$. The truthmakers of $X = 0 \vee Y = 0$ are the submodels $\mathcal{M}[X = 0]$, $\mathcal{M}[Y = 0]$ and $\mathcal{M}[X = 0 \wedge Y = 0]$. The first two submodels validate $D = max(X, Y) = 1$ since the second soldier is not affected by the intervention, and so $(X = 0 \vee Y = 0) \mathrel{\Box\!\!\rightarrow} D = 0$ is false at $\mathcal{M}$.

Briggs' extension of CMS allows us to assign a truth value to counterfactuals with disjunctive antecedents—and in fact, to counterfactuals with arbitrary Boolean compounds of atomic formulas in the antecedent. The main innovation to CMS consists in evaluating counterfactuals in the submodels that truthmake the antecedent. Implicit in Briggs' approach is a relevance principle for the truth conditions of counterfactuals, which we will also use later when defining their probability:

- **Relevance Principle (Truth Conditions)**. The truth value of a counterfactual $A \mathrel{\Box\!\!\rightarrow} B$ at a causal model $\mathcal{M}$ depends exclusively on the truth value of B in the submodels $\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_n$ generated by the interventions on the variables in $\mathcal{M}$ that truthmake A.

We now proceed to developing our proposal in the framework of probabilistic causal models.

5.1.3 Probabilistic Causal Modeling

In this section, we introduce probabilistic causal models and explain how CMS assigns a probability to counterfactuals. We will also see how the problem of the limited expressive power of CMS re-emerges at the probabilistic level: causal modeling semantics does not allow to assign a probability to counterfactuals with disjunctive antecedents.

Definition 5.7. *A probabilistic causal model is a tuple* $\mathcal{M} = \langle \mathcal{V}, \mathcal{G}, P \rangle$ *where*

- $\mathcal{V}$ *is a set of variables;*
- $\mathcal{G} \subset \mathcal{V} \times \mathcal{V}$ *is a set of directed edges between the variables in* $\mathcal{V}$*, defining the parents and descendants of each variable;*
- P *is a probability distribution on* $\mathcal{V}$ *subject to the Markov condition, that is, each variable* V *is probabilistically independent of its non-descendants, conditional on its parents.*

In a probabilistic causal model, the behavior of exogenous variables, and the dependencies of the endogenous variables on their parents, are described via a probability distribution. This differs from the non-probabilistic causal models in that, in the latter, variables are governed by structural equations.[3] Consider again the execution scenario with the probability distribution P described in Table 5.1. Thanks to the Markov condition, it is sufficient to specify the probability of the exogenous variables, and the conditional probability of the endogenous variables, given the values of their parents.

[3]The probabilistic nature of the models does not entail that the mechanism of dependence is intrinsically non-deterministic: for example, Pearl (2000, p. 26) seems to favor the view that the non-deterministic dependencies of the variables are due the lack of knowledge about the underlying deterministic mechanism.

C	
1	0.5
0	0.5

C	X	
	1	0
1	0.9	0.1
0	0.1	0.9

C	Y	
	1	0
1	0.9	0.1
0	0.1	0.9

X	Y	D	
		0	1
1	0	0.5	0.5
0	1	0.5	0.5
0	0	0.9	0.1
1	1	0.1	0.9

Table 5.1: Probability distribution for the variables in the execution example, as a function of the values of their parents. Intuitively, the table describes the dependencies among the variables; for instance we have that the value of X will be 1 with 90% probability if the value of C is 1, i.e. $P(X = 1|C = 1) = 0.9$. This means that it is almost certain that the executioner X fires under the order of the captain, but there is a little chance (10%) that X might miss the shot, (for example, if the weapon jammed). Also, according to the table, it is almost certain that the prisoner dies if both the executioners fire, $P(D = 1|X = 1, Y = 1)$, but there is a little chance (10%) that he might survive. Of course, this probability distribution must be intended as a toy example.

Analogously to the non-probabilistic case, probabilistic causal models provide an excellent tool for reasoning about counterfactuals. Again, the notion of an intervention is crucial. Pearl (2000) proposes that the probability of a counterfactual $A \boxright B$ at a probabilistic causal model $\mathcal{M}$, given a certain evidence E, amounts to the probability of B in the submodel generated by the intervention $do(A)$ after updating on E, where A is an atomic formula or a conjunction of atomic formulas. More formally, the following procedure describes ho to assign a probability to a counterfactual within a causal models:

Definition 5.8. *Consider a probabilistic causal model $\mathcal{M} = \langle \mathcal{V}, \mathcal{G}, P \rangle$ and a counterfactual $A \boxright B$ in the language $\mathcal{L}^{\boxright}_{CMS}$. Then the probability of $A \boxright B$, $P(A \boxright B)$, is computed according to the following procedure:*

1. *Update the probability $P(U = u)$ of each exogenous variable U on the evidence E, via Bayesian conditionalization, to the new probability $P'(U = u) = P(U = u|E)$, without changing the conditional dependencies among the variables. This is because the evidence should not change the structure of the causal relationships between the variables: it just informs us which context we are likely to be in (see Pearl 2000, pp. 33-38). So P' induces a new probability distribution on the (endogenous) variables, too.*

2. *Perform the intervention $do(A)$ on $\mathcal{M}$ to obtain a new submodel $\mathcal{M}'$ of $\mathcal{M}$; accordingly, change the probability distribution P' so that variables involved in the intervention do not depend on their parents anymore.*

3. *Use the new submodel* $\mathcal{M}' = \langle \mathcal{V}, \mathcal{G}', P'_{do(A)} \rangle$ *with post-intervention graph* $\mathcal{G}' \subseteq \mathcal{G}$ *and probability distribution* $P'_{do(A)}(\cdot)$ *to calculate the probability of* B *at* $\mathcal{M}'$ *(i.e.* $P'_{do(A)}(B)$*).*

At the end, the probability of $A \mathbin{\Box\!\!\rightarrow} B$*, after having learn the evidence* E*,* $P(A \mathbin{\Box\!\!\rightarrow} B \mid E)$ *amounts to:* $P'_{do(A)}(B)$

For example, consider the probabilistic execution model with the numbers from Table 5.1. Assume that we have learned about the death of the prisoner, without knowing whether the captain has given the signal, or whether the executioners have fired. We have thus learnt the evidence $E = \{D = 1\}$. By the procedure specified above, we need to update the probability of the *exogenous variables*, i.e., $P'(C = 1) = P(C = 1|D = 1) = 0.82$, which induces a new probability distribution P' on the endogenous variables.[4] Now, we want to compute the probability of $D = 0$ under the counterfactual assumption that X has not fired, $X = 0$, corresponding to the probability to the counterfactual "if executioner X hadn't fired, then the prisoner would not have died" ($X = 0 \mathbin{\Box\!\!\rightarrow} D = 0$). Following the above procedure, we should intervene by assigning value zero to X; this intervention $do(X = 0)$ can be understood as an external action that forces the prisoner not to fire, for instance we sabotage X's weapon. The action does not affect the probability of variables causally upstream of X: indeed our action is limited to X and does not influence the behavior of C. Instead, it preempts the causal power of C on X, and therefore we delete the arrow connecting C to X. However, this intervention does affect the variables causally *downstream* of X, imposing a new distribution on the model. Indeed, if we want how the prisoner is affected by this intervention, we need to calculate $P'_{do(X=0)}(D = 0)$. Following the above procedure, we obtain that

$$\begin{aligned} P'_{do(X=0)}(D=0) &= \sum_{y,c\in\{0,1\}} P(D=0|X=0,Y=y) \times P(Y=y|C=c) \times P(C=c|D=1) \\ &= 0.598. \end{aligned}$$

In other words, it is 59.8% probable that the prisoner would not have died under the counterfactual supposition that the executioner X hadn't fired. This is, by the way, much less than the conditional probability $P'(D = 0|X = 0) = 0.752$ because *updating on* $X = 0$ (with all other variables being unknown) would suggest an inference to the best explanation, i.e., that the captain did not give the signal. Hence, also the probability of $Y = 0$ goes up sharply when we learn $X = 0$, and so does the probability of $D = 0$.

[4] Henceforth, unless otherwise stated, we will use P' to refer to the probability distribution induced by $P'(C = 1) = P(C = 1|E) = 0.82$.

Like deterministic CMS, the probabilistic framework does not account for the probability of counterfactuals with disjunctive antecedents since interventions are only defined for atomic formulas and their conjunctions. We will now develop a proposal that expands probabilistic CMS to arbitrary Boolean compounds of atomic formulas in the antecedent, similar to what Briggs has achieved for deterministic CMS.

5.2 Expanding Causal Modeling to Complex Counterfactuals

In this section, we will propose a procedure to account for the probability of counterfactuals with disjunctive antecedents within a causal model. Our final result will follows from combining Briggs' (2012) work with the findings in (Eva, Stern, and Hartmann 2019).

5.2.1 CMS with Similarity Metrics

Suppose that we want to use probabilistic CMS in order to calculate the probability of a counterfactual with disjunctive antecedents. When we apply Pearl's procedure described in the previous section, steps 2 and 3 fail because the model generated by the intervention $do(X = 0 \vee Y = 0)$ is not well defined and consequently we cannot compute $P'(D = 0)$.

A first step toward solving this problem is to impose a probabilistic version of the **Relevance Principle** which was implicit in Briggs' framework:

- **Relevance Principle (Probability)**. The probability of a counterfactual $A \boxright B$ at a causal model $\mathcal{M}$ depends exclusively on the probability of B in the submodels $\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_n$ generated by the interventions on the variables in $\mathcal{M}$ that truthmake A.

Thus, the probability of $(X = 0 \vee Y = 0) \boxright D = 0$ depends exclusively on the probability of $D = 0$ in the three submodels generated by $do(X = 0)$, $do(Y = 0)$ and $do(X = 0 \wedge Y = 0)$. See Table 5.2. Step 2 is working now: performing the intervention $do(X = 0 \vee Y = 0)$ amounts to selecting three *specific* submodels. However, step 3 is still problematic: it is not clear how the probabilities of $D = 0$ in the three submodels should be combined. In fact, for $P'_{do(X=0)}(D = 0) = P'_{do(Y=0)}(D = 0) = 0.598$, whereas $P'_{do(X=0,Y=0)}(D = 0) = 0.9$.

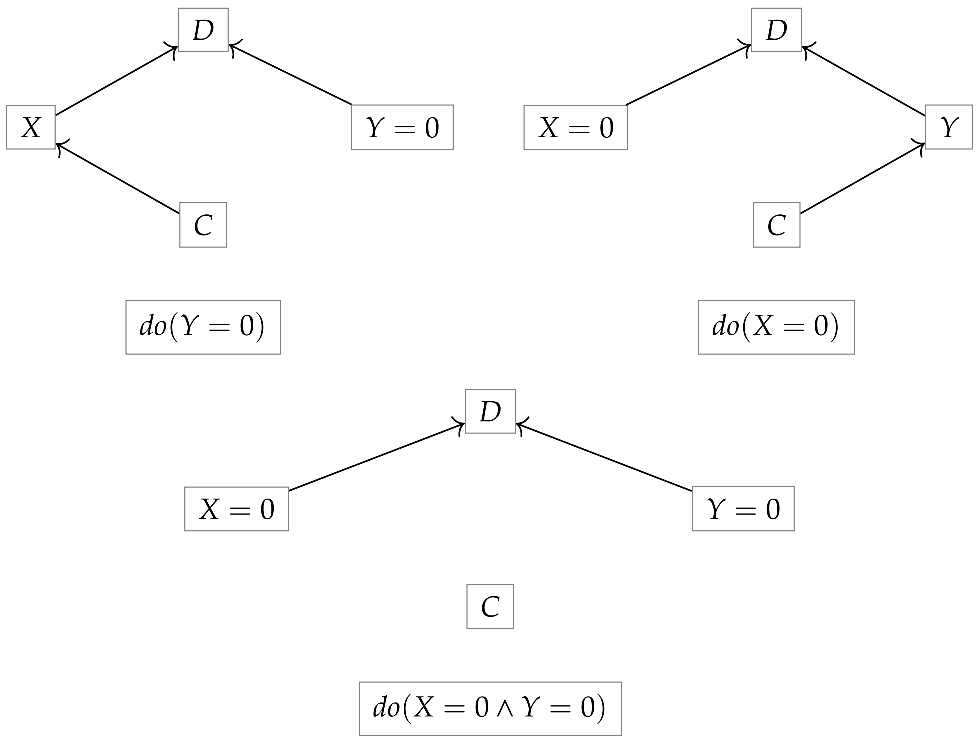


Table 5.2: The three submodels that truthmake the sentence $X = 0 \vee Y = 0$ in the execution example, with the interventions used to generate them.

It is clear that Briggs' solution for the truth conditions of a counterfactual with disjunctive antecedents will not help. There, the consequent needed to be true in *all* states that truthmake the antecedent. Briggs (2012, pp. 152-154) recognizes that this is a *choice*. The motivation is that there is no convincing argument for preferring a specific submodel. Moreover, also in Lewis' semantics, whenever there is a tie between the closest possible φ-worlds to a given one, i.e. $f(\varphi, w)$ contains more than one element, a counterfactual $\varphi \mathrel{\Box\!\!\rightarrow} \psi$ is evaluated as true only if ψ holds in *all* of these worlds, i.e. $f(\varphi, w) \subseteq [\psi]$. While this is a reasonable choice in the context of a *logic* of counterfactuals, we cannot transfer it to the *probability* of counterfactuals where the output of the submodels are not Boolean values, but real numbers. We need to assign *relative weights* to the truthmaking submodels, and this problem is specific to the probabilistic extension of Briggs' approach.

A natural requirement is that the values of $P'_s(B)$ in the relevant submodels indexed by s should *bound* the overall probability of the counterfactual $A \mathrel{\Box\!\!\rightarrow} B$ from above and below:

- **Convexity Principle**. For the probability of a counterfactual $A \mathrel{\Box\!\!\rightarrow} B$ at a probabilistic causal model $\mathcal{M}$, and the set of submodels $|A|_{\mathcal{M}}$ where we intervene on the variables in $\mathcal{M}$ as to truthmake A,

  $$\min(\{P_s(B) : s \in |A|_{\mathcal{M}}\}) \leq P(A \mathrel{\Box\!\!\rightarrow} B) \leq \max(\{P_s(B) : s \in |A|_{\mathcal{M}}\})$$

  where P_s denotes the probability distribution of the variables in submodel s, after updating on the available evidence and performing the truthmaking intervention.

In other words, the probability of a counterfactual cannot be greater (smaller) than the maximum (minimum) probability of the consequent in the causal models that truthmake the antecedent (see also Pearl 2017, p. 9).

The Convexity Principle still leaves space for a large class of weighting functions. A natural starting point is the *straight average* of $P'(D = 0)$ in the three submodels generated by $do(X = 0 \vee Y = 0)$. In this way, we would obtain $P'_{do(X=0 \vee Y=0)}(D = 0) = \frac{0.598+0.598+0.9}{3} = 0.698$. However, straight averaging is at best a default assumption and devoid of a compelling motivation. An alternative is to make the relative weight of the three submodels generated by $do(X = 0)$, $do(Y = 0)$ and $do(X = 0, Y = 0)$ depend on their degree of *similarity to the original model*. Once we have weights $\alpha_1, \alpha_2, \alpha_3$ for each of them, we can compute the post-intervention probability as

$$P'_{do(X=0 \vee Y=0)}(D=0) = \alpha_1 \times 0.598 + \alpha_2 \times 0.598 + \alpha_3 \times 0.9.$$

The question is how to measure this degree of similarity. A possible answer comes from a recent work of Eva, Stern, and Hartmann (2019) where the authors introduce two notions of similarity distance between causal models: *evidential* similarity distance, based on the shared probabilistic (in)depencies, and *counterfactual* similarity distance, based on shared counterfactual dependencies. In what follows, we restrict our attention to the latter since probabilistic independencies can hide true causal and counterfactual dependencies.[5]

Definition 5.9 (Eva, Stern, and Hartmann 2019). *Given a probabilistic causal model $\mathcal{M} = \langle \mathcal{V}, \mathcal{G}, P \rangle$, we can define the following notions:*

- ***Counterfactual Dependence between variables**. A variable V_2 is counterfactually dependent on another variable V_1 when an intervention on V_1 affects the probability distribution of V_2, i.e., for some $v \in \mathcal{R}(\mathcal{V}_1)$, $P_{do(V_1=v)}(V_2) \neq p(V_2)$.*[6]
- ***Counterfactual Similarity Distance** Two (probabilistic) causal models $\mathcal{M}$ and $\mathcal{M}'$ are more or less similar to each other, the more counterfactual dependencies they agree on. Specifically, the counterfactual distance between $\mathcal{M}$ and $\mathcal{M}'$ is the absolute value of the difference of their counterfactual dependencies normalized by the total number of possible counterfactual dependencies:*

$$d(\mathcal{M}, \mathcal{M}') = \frac{|C_{\mathcal{M}} - C_{\mathcal{M}'}|}{N_C} \in [0, 1].$$

Recall that a variable V_2 is counterfactually dependent on another variable V_1 if we can go from V_1 to V_2 by following a sequence of arrows from V_1 to V_2: arrows represent the structural equations, i.e., the *mechanisms* or *laws* that connect variables. Hence, if two models disagree on some counterfactual dependencies among the variables, they disagree on the *mechanism* connecting those variables. So, intuitively, the more laws governing the original model are broken in $\mathcal{M}'$, the more counterfactual-distant from $\mathcal{M}$ a causal model $\mathcal{M}'$ is (see also Lewis 1973a).

There are two principled options for calculating the probability of counterfactuals. First, we could focus on the submodel that is most similar to $\mathcal{M}$ in the above metric, and neglect the contribution of the other submodels. This is

[5] In the causal modeling literature, this is known as failure of the Faithfulness Condition.

[6] For example, in the execution model, D counterfactually depends on X, Y and C; while X and Y counterfactually depends on C.

feasible, but it would privilege a particular model and a specific way of truthmaking the antecedent. This is especially implausible when the truthmaking models have a similar distance to the original model and express qualitatively different ways of changing the mechanisms to make the antecedent true.

Second, we could propose that the weight of each submodel $\mathcal{M}'$ should be inversely proportional to its distance to the original model $\mathcal{M}$, according to the above distance measure. This is our preferred approach since it takes into account all relevant submodels that truthmake the antecedent (and only them).

For example, consider the execution story and the three submodels generated by $do(X = 0)$, $do(Y = 0)$ and $do(X = 0 \wedge Y = 0)$. The number of total pairwise counterfactual dependencies is $N_C = 12$; the original model $\mathcal{M}$ encodes $C_{\mathcal{M}} = 5$ counterfactual dependencies; each of the models generated by $do(X = 0)$ and $do(Y = 0)$ encodes $C_{\mathcal{M}'} = 4$ counterfactual dependencies and the model generated by $do(X = 0 \wedge Y = 0)$ encodes $C_{\mathcal{M}'} = 2$ counterfactual dependencies. Table 5.3 describes the counterfactual dependencies of the execution story and its submodels, where $V_1 \mathrel{\Box\!\!\rightarrow} V_2$ means that V_2 counterfactually depends on V_1:

	Original Model	$do(X=0)$	$do(Y=0)$	$do(X=0 \wedge Y=0)$
$C \mathrel{\Box\!\!\rightarrow} X$	Yes	No	Yes	No
$C \mathrel{\Box\!\!\rightarrow} Y$	Yes	Yes	No	No
$C \mathrel{\Box\!\!\rightarrow} D$	Yes	Yes	Yes	No
$X \mathrel{\Box\!\!\rightarrow} D$	Yes	Yes	Yes	Yes
$X \mathrel{\Box\!\!\rightarrow} Y$	No	No	No	No
$X \mathrel{\Box\!\!\rightarrow} C$	No	No	No	No
$Y \mathrel{\Box\!\!\rightarrow} D$	Yes	Yes	Yes	Yes
$Y \mathrel{\Box\!\!\rightarrow} X$	No	No	No	No
$Y \mathrel{\Box\!\!\rightarrow} C$	No	No	No	No
$D \mathrel{\Box\!\!\rightarrow} X$	No	No	No	No
$D \mathrel{\Box\!\!\rightarrow} Y$	No	No	No	No
$D \mathrel{\Box\!\!\rightarrow} C$	No	No	No	No

Table 5.3: Counterfactual Dependencies for the Execution Example.

Call $\mathcal{M}$ the original execution model. By looking at the table we can deduce that

$$d(\mathcal{M}, \mathcal{M}[X = 0]) = \frac{1}{12} \qquad d(\mathcal{M}, \mathcal{M}[Y = 0]) = \frac{1}{12}$$
$$d(\mathcal{M}, \mathcal{M}[X = 0 \wedge Y = 0]) = \frac{3}{12}$$

So, $\mathcal{M}[X=0]$ and $\mathcal{M}[Y=0]$ are equally similar to $\mathcal{M}$ and $\mathcal{M}[X=0 \wedge Y=0]$ is the most distant from $\mathcal{M}$. Hence, $\mathcal{M}[X=0 \wedge Y=0]$, which is the most distant submodel, will receive the least weight. Call $|A|_{\mathcal{M}} = \{s \mid s \Vdash A\}$ the set of truthmakers of A, i.e., the submodels generated by the intervention $do(A)$ on $\mathcal{M}$. In the model $\mathcal{M}$ of the execution story,

$$|X=0 \vee Y=0|_{\mathcal{M}} = \{\mathcal{M}[X=0], \mathcal{M}[X=0], \mathcal{M}[X=0 \wedge Y=0]\}.$$

For $s \in |X=0 \vee Y=0|_{\mathcal{M}}$, we define its weight as

$$\alpha(s) = \frac{d(\mathcal{M},s)^{-1}}{\sum_{t \in |X=0 \vee Y=0|_{\mathcal{M}}} d(\mathcal{M},t)^{-1}},$$

following the rationale that the weight should be inversely proportional to the distance from the original model, normalized by the sum of all weights.

By some computation, we get that

$$\alpha(\mathcal{M}[X=0]) = \alpha(\mathcal{M}[Y=0]) = \frac{3}{7} \qquad \alpha(\mathcal{M}[Y=0 \wedge X=0]) = \frac{1}{7}$$

Applied to the execution story, we then find that

$$P'((X=0 \vee Y=0) \mathrel{\Box\!\!\rightarrow} D=0) = \frac{3}{7} \times 0.598 + \frac{3}{7} \times 0.598 + \frac{1}{7} \times 0.9 \approx 0.64,$$

in agreement with the Convexity Principle. We can generalize the weighting procedure as follows:

Definition 5.10. *for a causal model $\mathcal{M}$, for an arbitrary formula $A \in For_{\mathcal{L}_{CMS}}$, for $s \in |A|_{\mathcal{M}}$,*

$$\alpha(s) = \frac{d(\mathcal{M},s)^{-1}}{\sum\limits_{t \in |A|_{\mathcal{M}}} d(\mathcal{M},t)^{-1}}.$$

Consequently, we calculate the probability of a counterfactual $A \mathrel{\Box\!\!\rightarrow} B$ in our language $\mathcal{L}^{\Box\!\rightarrow}_{CMS}$ relative to a causal model $\mathcal{M}$, as

$$\begin{aligned} P(A \mathrel{\Box\!\!\rightarrow} B) &= \sum_{s \in |A|_{\mathcal{M}}} \alpha(s) \times p_s(B) \\ &= \sum_{s \in |A|_{\mathcal{M}}} \frac{d(\mathcal{M},s)^{-1}}{\sum\limits_{t \in |A|_{\mathcal{M}}} d(\mathcal{M},t)^{-1}} \times p_s(B) \end{aligned} \tag{5.1}$$

Equation (5.1) expresses our main idea in a nutshell: the probability of the counterfactual $P(A \mathrel{\Box\!\!\rightarrow} B)$ is the probability of the consequent B in all submodels that truthmake the antecedent, weighted inversely by their similarity to the original model, where similarity is measured by the number of shared counterfactual dependencies. Our account thus synthesizes Causal Modeling Semantics with the Relevance Principle (=focusing on models that truthmake the antecedent, as in (Briggs 2012)) and (Eva, Stern, and Hartmann 2019) proposal for measuring similarity between causal models.

It is easy to see that our definition of the probability of a counterfactual with disjunctive antecedents extends to more complex sentences, too. Fine's truthmaker semantics indicates the truthmaking space states of all Boolean compunds of atomic sentences. Thus, for any sentence that we wish to take as the antecedent of a counterfactual, we simply determine the truthmaking states, the interventions on the causal model that correspond to them, and the corresponding counterfactual probabilities. Then we can use the Eva-Stern-Hartmann procedure for weighting the causal models that correspond to the truthmaking states.

For example, if, for binary variables A and B, our counterfactual is "if $A = B$, then $C = 1$" (with actual values $A = 1$ and $B = 0$), the antecedent has two truthmakers: the model generated by $do(A = 1, B = 1)$ and the one generated by $do(A = 0, B = 0)$. The two causal models obtained will then have the same weight according to our procedure, since the intervention affects the same variables and yields the same counterfactual dependencies. In other words, the probability of the counterfactual "if $A = B$, then $C = 1$" is simply the straight average of the probability of $C = 1$ under the interventions $do(A = 1, B = 1)$ and $do(A = 0, B = 0)$.[7]

Taking stock, we have developed a procedure that goes beyond the achievements of Galles and Pearl (1998) and Halpern (2000), who can calculate probabilities of counterfactuals, but only for antecedents representing a (conjunctive) set of interventions. On the other hand, Briggs (2012) has a general logic of counterfactuals, allowing for arbitary Boolean compounds as antecedents, but no extension to probabilistic reasoning. Our contribution provides a probabilistic counterpart of her logic motivated from the very same principles.

5.2.2 Back to Lewis: Comparison with Imaging

In this section, we compare our account to the predictions of generalized imaging (Gärdenfors 1982; Lewis 1976) for assigning a probability to a counterfactual. Specifically, we will focus on a particular kind of generalized

[7] Note that this also holds if it is *actually* the case that $A = B = 1$. Calculating the probability of the counterfactual does not privilege the actual values of variables; all that matters is whether the distance of the truthmaking models from the original model in terms of counterfactual dependencies.

imaging, called *Bayesianized imaging* by Joyce (1999). According to our notation in Definition 4.2, Bayesianized imaging can be defined as follows:

Definition 5.11. *Given a functional Lewisian model $\mathcal{M} = \langle W, f, \vDash \rangle$ satisfying* **Centering**, *consider a probability distribution P over W. Moreover, assume that for all $\varphi \in For_{\mathcal{L}_{\Box\!\rightarrow}}$ such that $\nvDash_{\mathbf{VC}} \neg\varphi$, for all $w \in W$, $f(\varphi, w) \neq \emptyset$.*

Consider a formula $\varphi \in For_{\mathcal{L}_{\Box\!\rightarrow}}$, and a function $T_\varphi : W \times W \to [0,1]$ defined as follows:

$$T_\varphi(w, w') = \frac{P(w)}{\sum\limits_{w' \in f(\varphi, v)} P(w')}$$

Then, the bayesianized imaged probability distribution P^φ is defined as follows:

$$P^\varphi(w) = \sum_{v \in W} P(v) \times \begin{cases} \frac{P(w)}{\sum_{w' \in f(\varphi,v)} P(w')} & \text{if } w \in f(\varphi, v) \\ 0 & \text{otherwise} \end{cases} \tag{5.2}$$

In this case, each world w where φ is false redistributes its weight among the closest worlds where φ is true, in proportion to the prior probability of these worlds.

There is a deep connection between Bayesianized imaging and CMS. Pearl (2017) shows that the probability of a counterfactual $A \Box\!\!\rightarrow B$, in the language $\mathcal{L}^{\Box\!\rightarrow}_{CMS}$ with $A = A_1 \wedge .... \wedge A_n$ being a conjunction of atomic formulas, can be characterized in two equivalent ways: either, in Causal Modeling Semantics, by

$$P(A \Box\!\!\rightarrow B) := P_{do(A)}(B) \tag{5.3}$$

or, *when we count worlds with equal causal histories as equally similar*, and use the Bayesianized imaging function P^A from Equation (5.2), by

$$P(A \Box\!\!\rightarrow B) := P^A(B) = \sum_{w \vDash B} P^A(w) \tag{5.4}$$

The first condition ("equal causal history") means that the most similar A-worlds to a $\neg A$-world w contain all and only those A-worlds that agree with w on the value of the variables that cannot be affected by $do(A)$, i.e., the non-descendants of A. In this settings, possible worlds coincide with complete configurations of variables For example, in the execution model, a possible world will be $w = \langle C = 1, X = 1, Y = 1, D = 1 \rangle$.

Pearl then shows that these two characterizations are equivalent, i.e.,

$$P^A(B) = P_{do(A)}(B). \tag{5.5}$$

In other words, the transformation defined by the *do*-operator can, for atomic interventions or their conjunctions, be interpreted as an imaging-type mass-transfer. This is a significant result showing that Bayesianized imaging and CMS agree for a large class of interventions. This result also motivates why we put Bayesianized imaging (as opposed to, e.g., equal weights imaging) at the center of the comparison of our own proposal with SLSS.

Worlds	Values				Closest worlds for imaging w_i on $X = 0 \vee Y = 0$	
	C	X	Y	D	Option 1: $f_1(w_i) = \ldots$	Option 2: $f_2(w_i) = \ldots$
w_1	1	1	1	1	$\{w_3, w_4, w_7, w_8\}$	$\{w_3, w_4, w_5, w_6, w_7, w_8\}$
w_2	1	1	1	0	$\{w_3, w_4, w_7, w_8\}$	$\{w_3, w_4, w_5, w_6, w_7, w_8\}$
w_3	1	1	0	1	$\{w_3\}$	$\{w_3\}$
w_4	1	1	0	0	$\{w_4\}$	$\{w_4\}$
w_5	1	0	1	1	$\{w_5\}$	$\{w_5\}$
w_6	1	0	1	0	$\{w_6\}$	$\{w_6\}$
w_7	1	0	0	1	$\{w_7\}$	$\{w_7\}$
w_8	1	0	0	0	$\{w_8\}$	$\{w_8\}$
w_9	0	1	1	1	$\{w_{11}, w_{12}, w_{15}, w_{16}\}$	$\{w_{11}, w_{12}, w_{13}, w_{14} w_{15}, w_{16}\}$
w_{10}	0	1	1	0	$\{w_{11}, w_{12}, w_{15}, w_{16}\}$	$\{w_{11}, w_{12}, w_{13}, w_{14} w_{15}, w_{16}\}$
w_{11}	0	1	0	1	$\{w_{11}\}$	$\{w_{11}\}$
w_{12}	0	1	0	0	$\{w_{12}\}$	$\{w_{12}\}$
w_{13}	0	0	1	1	$\{w_{13}\}$	$\{w_{13}\}$
w_{14}	0	0	1	0	$\{w_{14}\}$	$\{w_{14}\}$
w_{15}	0	0	0	1	$\{w_{15}\}$	$\{w_{15}\}$
w_{16}	0	0	0	0	$\{w_{16}\}$	$\{w_{16}\}$

Table 5.4: Two plausible selection functions f_1 and f_2 in the execution example with disjunctive interventions. The two selection functions correspond to two different ways of identifying, for any $w_i \in W$, the closest possible world where $X = 0 \vee Y = 0$ holds.

In principle, we can extend Bayesianized imaging to the probability of counterfactuals with disjunctive antecedents. Consider the execution model again. We associate a possible world w to each possible realization of the binary variables C, X, Y, D; so there are 16 possible worlds in total. The probability of each of them is simply the joint probability of the realizations of the variables in that possible world, respecting the conditional independence relations imposed by model $\mathcal{M}$ and the Causal Markov Condition. For modeling Bayesianized imaging on a sentence A, we develop a three-step procedure analogous to the one recommended by CMS:

Definition 5.12. *Given a probabilistic causal model $\mathcal{M} = \langle \mathcal{V}, \mathcal{G}, P \rangle$, and a counterfactual $A \mathrel{\Box\!\!\to} B$, we can compute the probability of $A \mathrel{\Box\!\!\to} B$ according to the following procedure:*

1. *Update the prior probability of the exogenous variables U on the observed evidence E from $P(U = u)$ to the posterior probability $P'(U = u) = P(U = u|E)$. For all endogenous variables, their conditional probability distribution continues to be given by the probabilistic causal model $\mathcal{M}$.*

2. *Transfer the mass of the $\neg A$-worlds to the closest possible A-worlds (chosen by the selection function f), weighted by the posterior probability of the latter. This will yield the probability function $P'^{A}(\cdot)$.*

3. *Calculate the probability of any sentence B as $P'^{A}(B)$.*

In the execution model, only C is an exogenous variable and this means that the joint posterior distribution after the first step of the above procedure will look as follows:

$$P'(C, X, Y, D) = P'(C) \times P(X|C) \times P(Y|C) \times P(D|X, Y)$$

Now we proceed to the second step and image P' on $(X = 0 \vee Y = 0)$. This means that four worlds will have weight zero in $P'_{X=0 \vee Y=0}$: w_1, w_2, w_9 and w_{10} in Table 5.4. The question is how their weight should be distributed to the rest; and this depends on what are the closest neighbors to these possible worlds.

The first conceptual obstacle in defining a similarity order is to decide which variables are *not* affected by $do(X = 0 \vee Y = 0)$. Again, we translate the problem into Causal Modeling Semantics. According to Briggs (2012), the disjunctive intervention $do(X = 0 \vee Y = 0)$ can be regarded as encoding three different interventions, $do(X = 0)$, $do(Y = 0)$, and $do(X = 0 \wedge Y = 0)$. The closest worlds to w_1 for the first intervention are w_7 and w_8, for the second, they are w_3 and w_4, and for the third, w_5 and w_6. Depending on how seriously we consider the option of intervening on *both* variables as a way of expressing $do(X = 0 \vee Y = 0)$, this gives us two options for the most similar worlds to w_1: $\{w_3, w_4, w_7, w_8\}$ or $\{w_3, w_4, w_5, w_6, w_7, w_8\}$. And vice versa for the other worlds whose weight needs to be canceled. Both options are represented in the rightmost columns of Table 5.4.[8]

However, if we calculate the probability of the counterfactual $(X = 0 \vee Y = 0) \mathrel{\Box\!\!\rightarrow} D = 0$, after having learnt the evidence $D = 1$, the result of Bayesianized imaging will, for either of these similarity orders, differ from our proposal. For Option 1, we obtain $P'^{X=0 \vee Y=0}(D = 0) \approx 0.56$, and for Option 2, we obtain $P'^{X=0 \vee Y=0}(D = 0) \approx 0.57$.[9] This is arguably a bad prediction

[8] A potential third option that also takes into account the value of D, i.e., $f(w_1) = \{w_3, w_7\}$, does not yield qualitatively different results.

[9] Alessandro Zangrandi's GitHub https://github.com/zazangra/lewis_imaging offers a Python program to perform Bayesianized imaging on a causal model.

since it violates the plausible Convexity Principle: the probability of the counterfactual should be bounded from above and below by the (maximal and minimal) probability of the consequent in the causal submodels that truthmake the antecedent. To recall:

$$P'(X = 0 \;\Box\!\!\rightarrow D = 0) = 0.598 \qquad P'((X = 0 \wedge Y = 0) \;\Box\!\!\rightarrow D = 0) = 0.9$$
$$P'(Y = 0 \;\Box\!\!\rightarrow D = 0) = 0.598$$

To the extent that the Convexity Principle is plausible and compelling, we should reject any procedure that violates this constraint. Why should the probability of the counterfactual be above or below the probability of the consequent in all relevant submodels? It is simply paradoxical that the death of the prisoner, $D = 1$, is more probable under the hypothetical assumption that *at least one* of the two executioners did not fire ($P'^{X=0 \vee Y=0}(D = 0) \approx 0.56$) than under the assumption that *only one* did not fire ($P'^{X=0}(D = 0) = 0.598$).

Primarily, the failure of Convexity in imaging is due to the fact that there is no systematic connection between $P'^{X=0}(D = 0)$ and $P'^{X=0 \vee Y=0}(D = 0)$, like in our own proposal. For instance, when imaging on $X = 0$, part of the mass of w_3 is transferred to w_5, whose probability mass makes a contribution to $P'^{X=0}(D = 0)$, but not to $P'^{X=0 \vee Y=0}(D = 0)$ (in Option 1). This explains why the latter probability falls below $P'^{X=0}(D = 0)$, i.e., below the bounds resulting from the Convexity Principle. In other words, the violation of the Convexity Principle is due to the fact that Bayesianized imaging does not respect the Relevance Principle: the possible worlds do not contain any information about the causal structure of the model.

Of course, generalized imaging offers an entire universe of different mass transfer functions. So we do not exclude that the imaging theorist can find a function that complies with the Convexity Principle.[10] However, this must come at the price of choosing a procedure that deviates systematically from CMS for (conjunctions of) atomic interventions. What the imaging theorist *cannot* have is a probability mass transfer function that agrees in regular circumstances with CMS, and that satisfies at the same time the Convexity Principle when applied to more complex interventions. Indeed, Pearl (2017, pp. 6-7) explicitly advises caution when applying imaging to disjunctive interventions, such as the ones that we discussed in this paper. Hence, we conclude that the combination of Lewis' semantic idea and imaging has not yet delivered a convincing response to the problem of evaluating the proba-

[10]Equal weights imaging, a possible alternative, respects the Convexity Principle because it trivializes the problem: imaging on $X = 0$, $Y = 0$, $X = 0 \wedge Y = 0$ and $X = 0 \vee Y = 0$ all yield the same probability $P'(D = 0) = 0.5615$. This is obviously an unacceptable result.

bility of counterfactuals with disjunctive antecedents within the framework of causal modeling.

5.3 Conclusions

In this chapter, we have extended Causal Modeling Semantics to the evaluation of the probability of counterfactuals with disjunctive antecedents, and more generally, to any counterfactuals whose antecedents are truth-functional compounds of atomic sentences. To the best of our knowledge, no other proposal has been advanced in the literature to achieve this goal. Our approach is very natural combines three well-established ideas: (1) Briggs' characterization of disjunctive interventions relying on truthmaking causal submodels; (2) weighting the contributions of these submodels according to their similarity with the original world; (3) Eva et al.'s definition of a similarity metric between causal models by counting shared counterfactual dependencies.

As an alternative to our approach, one can assign probabilities to counterfactuals with disjunctive antecedents by imaging mass transfers, and Bayesianized imaging in particular. However, we showed that this option does not return plausible predictions about the probability of counterfactuals. What is more, it violates intuitive requirements such as the Convexity Principle and the Relevance Principle taht we propose as intuitive and compelling constraints over an account of the probability of counterfactuals within causal modeling.

Concluding Remarks

The present thesis provides a novel perspective on Lewis counterfactuals and variably strict conditionals. Throughout this work, we have addressed several key questions that have previously been overlooked in the existing literature. Our focus has been on investigating the truth-conditions, logic, and probability aspects of Lewis variably strict conditionals. To tackle these inquiries, we have integrated methodologies from various disciplines, such as algebraic logic, Dempster-Shafer Theory, formal epistemology, artificial intelligence, and formal semantics. This interdisciplinary approach has allowed us to shed new light on a classical theme in logic and philosophy.

Nevertheless, our research has also uncovered potential avenues for further exploration. In the subsequent sections, we will highlight some areas that warrant further investigation. These points can pave the way for future research:

The Algebra of Counterfactuals

In Chapter 1, we have undertaken a systematic algebraic treatment of Lewis' variably strict conditional logics by introducing a clear distinction between the global and local companions of each variably strict conditional logic. Subsequently, we delved into analyzing their respective algebraic semantics. Despite this progress, several open problems remain in this area that warrant further investigation.

One essential area for exploration is the structural analysis of Lewis algebras to study their corresponding variety. To achieve this, it is crucial to characterize the deductive filters of global variably strict conditional logics over the corresponding Lewis algebras. This characterization will provide valuable insights into the properties and behavior of these algebras.

In a different direction, a duality theory connecting Lewis algebras with their dual topological structures would significantly contribute to a better understanding of the relationship between variably strict conditional logics and their possible worlds semantics. Such a duality theory would shed light on the nature of the equivalence between functional and spherical Lewisian models and help clarify the logical status of the limit assumption.

From a broader perspective, our work has opened the door for a more systematic analysis of the varieties of Lewis algebras and their connection with modal algebras. This exploration has the potential to yield valuable knowledge and insights into the underlying structures of these algebras, contributing to the advancement of the field.

Boolean Algebras of Counterfactuals

In Chapter 2, we have showed how, within the novel framework of Boolean algebras of conditionals, Lewis counterfactuals can be reduced to a combination of a modal and a conditional operator $(\cdot \mid \cdot)$ that resembles Adams conditional. Our findings not only establish a foundation for future research on Boolean algebras of conditionals and related topics, but also raise some intriguing questions that warrant further exploration.

Specifically, our results from Chapter 2 are limited to a simple language where nested occurrences of the counterfactual arrow are not allowed. Hence, one key open problem is to determine whether our results can also be extended to a more expressive language that allows for nested occurrences of the conditional operator, such as $\Box((\varphi \mid \psi) \mid \top)$. In principle, this extension seems possible as the BAC construction can be iterated multiple times, for instance, we could apply the BAC construction on a BAC, so to obtain $\mathcal{C}(\mathcal{C}(\mathbf{A}))$. However, investigating the properties preserved by this iteration and the structural features of nested BACs requires closer attention.

From a technical standpoint, a deeper logical investigation of the modal BACs framework presents an interesting avenue for research. Specifically, the logical consequence $\models_{LBC_\Box}$ has not yet been axiomatized, and developing a sound and complete logical system with respect to $\models_{LBC_\Box}$ would enhance our understanding of the logical behavior of expressions like $\Box(\varphi \mid \psi)$. In this thesis, we mainly focused on a specific kind of modal BACs, that is Lewis algebras. However, exploring various types of modal BACs based on different axioms imposed on the modal operators opens up new possibilities. The resulting modal BACs could induce novel logical consequence relations that may be stronger of weaker than $\models_{LBC_\Box}$. This research direction intersects with recent programs in proof theory. In particular, Girlando, Negri, and Olivetti (2021) have recently investigated stronger logics than variably strict conditional logics resulting from dropping the axiom (NE) from each variably strict conditional logic. It is plausible to conjecture that, for a fragment of their language, these logics correspond to some logical consequence relations definable over modal BACs by dropping the condition $(L3)$. Investigating these logical consequence relations and proving their soundness and completeness with respect to the logics introduced by Girlando, Negri, and Olivetti (2021) represents an interesting open problem that requires further investigation.

From a philosophical perspective, there are numerous intriguing directions to explore and assess the impact of BACs. Firstly, certain open questions remain regarding the interpretation of the modal operator involved in the definition of a counterfactual, i.e., $\Box(\varphi \mid \psi)$. While we have proposed some plausible suggestions, we believe there are many other viable interpretations to be considered. For example, exploring a deontic interpretation of $\Box$ as a *must* modality appears promising.

On a more foundational level, it would be valuable to further investigate the connections between BACs and the triviality results. Specifically, exploring whether BACs can offer insights into the nature of the proposition expressed by a probabilistic conditional is of great interest. We have observed how the BACs framework successfully establishes an equation between the proposition expressed by a conditional and the corresponding conditional probability. However, understanding how to interpret the proposition expressed by a conditional inside BACs requires clarification, as it cannot be simply identified with the set of worlds where that proposition holds true.

This line of research also raises questions about the relationships between the BACs, the logic $\vdash_{\mathbf{LBC}}$, and Adams' conditional logic. It is evident that the logic $\vdash_{\mathbf{LBC}}$ coincides with Adams' (1975) conditional logic, and they also share a common linguistic domain where nested occurrences of the conditional operator are not allowed. However, $\vdash_{\mathbf{LBC}}$ and Adams' logic arise from two different semantic accounts: $\vdash_{\mathbf{LBC}}$ is interpreted within the BACs framework, whereas Adams' logic posses a probabilistic semantics. Exploring the relationship between Adams' probabilistic semantics and the BACs semantics would provide valuable insights into how probabilistic semantics relates to a more standard truth-conditional semantics. Additionally, investigating whether BACs can serve as an algebraic foundation for Adams' probabilistic semantics would be an important contribution to this field.

In summary, the philosophical exploration of BACs opens up numerous possibilities. Addressing questions concerning the interpretation of the modal operator, the connection to triviality results, the relationship with Adams' conditional logic, and the role of BACs in probabilistic semantics will enrich our understanding of these logical structures and their broader implications. By delving into these areas, we can further advance the understanding and application of BACs in both philosophical and logical contexts within the framework of counterfactual reasoning.

Counterfactuals as Definable Conditionals

In Chapter 3, we have built upon the foundational work of van Fraassen (1974) and expanded on the results concerning modal BACs. Specifically, we demonstrated that Lewis variably strict conditional logics can be entirely translated into the corresponding **KV**-logical consequence. This translation was induced by a new type of possible worlds models called spherical Kripke models, which offer greater expressiveness compared to Lewisian models. Spherical Kripke models can effectively account for both Stalnaker conditionals and Lewis variably strict conditionals at the same object language level. Furthermore, we explored and axiomatized some logical consequences definable over these models, i.e. **KV**-logics, which closely mirror the corresponding variably strict conditional logics.

The framework of spherical Kripke models and their associated logic is new and deserves further investigation from both technical and philosophical perspectives. From a technical standpoint, it would be insightful to analyze the different classes of spherical Kripke models that can be defined. In this thesis, we focused solely on the class of spherical Kripke models that mirror variably strict conditional logics. However, other constraints can be imposed on the accessibility relations or the underlying systems of spheres. Understanding what kind of logics would result from imposing standard normal modal logic constraints on the accessibility relation is an intriguing question. It is plausible that standard properties of the accessibility relation may also impact the properties of the counterfactual conditional defined as $\Box(\varphi > \psi)$.

Moreover, exploring whether spherical Kripke models can serve as a reductionist semantic framework for other conditional operators is worth investigating. For example, by dropping the **Sphericality** condition from spherical Kripke models, it may be possible to translate the conditionals in the logics introduced by Girlando, Negri, and Olivetti (2021) into the corresponding $\Box(\varphi > \psi)$ within our language. This observation suggests that spherical Kripke frames and their associated logical consequences could potentially serve as translations for various conditional logics. However, the extent to which this is possible deserves further investigation.

Furthermore, we have noted that our newly introduced **KV**-logics are not closed under uniform substitution, which places them within the broader and recent logical framework of weak logics. Weak logics have garnered attention recently, with Nakov and Quadrellaro (2022) proposing a general algebraic treatment for these logics. Exploring whether and how these algebraic methods can be applied to our **KV**-logics is an interesting point for future research, as it may reveal peculiar properties of **KV**-logics.

From a philosophical perspective, the use of spherical Kripke models and **KV**-logics has allowed for a new reductionist account of variably strict conditionals, reinforcing the view that counterfactuals (and variably strict conditionals) can be interpreted in terms of a modal operator combined with a Stalnaker conditional. It would be intriguing to investigate whether this reductionist view can be extended to other types of conditionals, and consequently, whether spherical Kripke models can provide a unified account for different classes of conditionals as well. Furthermore, our findings raise natural questions regarding the correct interpretation of the modal operator involved in the newly proposed truth conditions for counterfactuals, i.e. $\Box(\varphi > \psi)$.

On a different note, the new interpretation of Lewis counterfactuals may have various philosophical implications. For instance, Lewis (1973a) argues that causation can be explained in terms of counterfactual dependence, where A is a cause of B if and only if both the counterfactuals $A \mathrel{\Box\!\!\rightarrow} B$ and $\neg A \mathrel{\Box\!\!\rightarrow} \neg B$ are true. This idea of reducing causal relationships to counterfactual dependence is also supported by Pearl (2000). Therefore, it would be interesting to explore, conceptually, whether our interpretation of a counterfactual as $\Box(\varphi > \psi)$ also impacts the interpretation of causal relationships. If we accept the view that causation can be reduced to counterfactual dependence, then our interpretation of counterfactuals suggests that causation can be reduced to a modality of a specific conditional dependence. However, further investigation into the relationship between causation and modality, and how our translations relate to the ongoing debate over causation, is required.

Probability of Counterfactuals

In Chapter 4, we demonstrated how belief functions can be utilized to characterize the probability of counterfactuals (and variably strict conditionals in general), thereby presenting a new probabilistic interpretation of Lewis conditionals. However, a precise and thorough exploration of the connections between variably strict conditional logics and belief functions is still required.

While we provided basic examples of how logical axioms transfer to constraints over the corresponding imaged belief functions, a more comprehensive and precise understanding of the connections between these two levels is still lacking. Establishing this connection would enable us to determine, for each axiom of variably strict conditional logic, the corresponding constraint on the induced imaged belief function. This knowledge would not only contribute to a deeper probabilistic interpretation of variably strict conditional logics, but also help characterize the logical counterpart of the imaging procedure for updating belief functions.

Another open question pertains to the relationship between imaged belief functions and conditional belief functions. Various proposals have been made regarding how conditionalization should operate within the context of Dempster-Shafer theory (see Coletti and Mastroleo 2006), but the specific connections between these proposals and the imaged belief functions we reviewed in this thesis remain unclear.

In Chapter 5, we expanded causal modeling semantics by introducing a newly defined procedure to compute the probability of counterfactuals with complex antecedents. From the perspective of causal modeling, an open philosophical question arises: are our findings in Chapter 5 truly an explication of the "probability of counterfactuals"? While our procedure provides results that differ from the characterization of Lewis counterfactuals in terms of belief functions, it remains uncertain whether these results correspond to the probability, plausibility, or assertability of the proposition expressed by the counterfactual sentence. In this respect, It would be illuminating to investigate whether our procedure can also be employed to characterize the probability of some conditionals within possible worlds models.

On the experimental side, our work on Causal Modeling Semantics (CMS) lends itself to testing. Designing linguistic experiments to assess the pragmatic plausibility of our convexity principle, which asserts that the probability of a counterfactual $A \mathrel{\Box\!\!\to} B$ should be bounded from above and below by the *best* and *worst* scenarios for B under the supposition of A, would provide valuable insights.

Another potential application of our work on CMS is to shed new light on the notion of disjunctive causes introduced by Sartorio (2006). By leveraging our newly introduced method, we may gain a deeper understanding of disjunctive causes and their implications within the framework of CMS.

Finally, it is essential to note that CMS and our newly introduced method are currently limited to a fragment of the language where nested occurrences of the counterfactual arrow are not allowed. Exploring the extension of our results to a more expressive language that permits nested counterfactuals presents an open problem that requires further investigation. Such an extension would significantly broaden the applicability and potential impact of CMS and our research.

In conclusion, our thesis has advanced the understanding of Lewis counterfactuals and variably strict conditionals by addressing critical aspects that were previously overlooked in the existing literature. By integrating diverse methodologies, we have demonstrated the fruitful application of these approaches to classical themes in logic and philosophy. Through our investigation, we have shed light on the truth-conditions, the logic, and the probability of Lewis variably strict conditionals. The use of algebraic logic, Dempster-Shafer Theory, formal epistemology tools, and formal semantics has provided valuable insights into these topics. Furthermore, our research has paved the way for future directions of inquiry that will contribute to the ongoing development of conditional logic and its philosophy. The points discussed above represent promising avenues for further exploration and expansion of the field.

Appendix A

Appendix of Chapter 0

A.1 Proof of Lemma 0.4

Proof. Consider a spherical Lewisian model $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$, we need to show that $\mathcal{M}^f = \langle W^f, f, \vDash^f \rangle$ is a functional Lewisian model according to Definition 0.3.

(i) By induction on the complexity of a formula, and by employing Lemma 0.1 and the definition of f in $\mathcal{M}^f$, it is immediate to see that $[\varphi]^f = [\varphi]$ for all $\varphi \in For_{\mathcal{L}_{\Box\!\!\rightarrow}}$.

(ii) It remains to show that f satisfies the relevant constraints.

1. for all $\varphi \in \mathcal{L}_{\Box\!\!\rightarrow}$, for all $w \in W$, $f(\varphi, w) \subseteq [\varphi]^f$. Consider any $v \in f(\varphi, w)$; by definition $v \in [\varphi] \cap min^{\varphi}_{\subseteq}(\mathcal{S}(w))$. Hence, clearly, $v \in [\varphi]$, and so, by (i), $v \in [\varphi]^f$;
2. for all $\varphi, \psi \in \mathcal{L}_{\Box\!\!\rightarrow}$, for all $w \in W$, if $f(\varphi, w) \subseteq [\psi]^f$ and $f(\psi, w) \subseteq [\psi]^f$, then $f(\varphi, w) = f(\psi, w)$. Assume $f(\varphi, w) \subseteq [\psi]^f$ and $f(\psi, w) \subseteq [\psi]^f$; this means that, by definition of f and by (i) above, $[\varphi] \cap min^{\varphi}_{\subseteq}(\mathcal{S}(w)) \subseteq [\psi]$ and $[\psi] \cap min^{\psi}_{\subseteq}(\mathcal{S}(w)) \subseteq [\varphi]$.

   We are going to show that $[\psi] \cap min^{\psi}_{\subseteq}(\mathcal{S}(w)) = [\varphi] \cap min^{\varphi}_{\subseteq}(\mathcal{S}(w))$. First, observe that if $min^{\varphi}_{\subseteq}(\mathcal{S}(w)) = min^{\psi}_{\subseteq}(\mathcal{S}(w))$, then, since by assumption $[\varphi] \cap min^{\varphi}_{\subseteq}(\mathcal{S}(w)) \subseteq [\psi]$ and $[\psi] \cap min^{\psi}_{\subseteq}(\mathcal{S}(w)) \subseteq [\varphi]$, we have that $[\psi] \cap min^{\psi}_{\subseteq}(\mathcal{S}(w)) = [\varphi] \cap min^{\varphi}_{\subseteq}(\mathcal{S}(w))$. If $min^{\varphi}_{\subseteq}(\mathcal{S}(w)) \neq min^{\psi}_{\subseteq}(\mathcal{S}(w))$, we reason as follows.

   $(\subseteq)$ Assume for reductio that there is a $v \in [\psi] \cap min^{\psi}_{\subseteq}(\mathcal{S}(w))$ such that $v \notin [\varphi] \cap min^{\varphi}_{\subseteq}(\mathcal{S}(w))$. Then, we have different cases to consider:

(a) $[\varphi] \cap min_{\subseteq}^{\varphi}(\mathcal{S}(w)) = \emptyset$. This means that $min_{\subseteq}^{\varphi}(\mathcal{S}(w)) = \emptyset$. However, by assumption $[\psi] \cap min_{\subseteq}^{\psi}(\mathcal{S}(w)) \subseteq [\varphi]$ and $v \in [\psi] \cap min_{\subseteq}^{\psi}(\mathcal{S}(w))$, hence $[\varphi] \cap \bigcup \mathcal{S}(w) \neq \emptyset$. So, by the limit assumption, there must be a minimal φ-permitting sphere in $\mathcal{S}(w)$, and so $min_{\subseteq}^{\varphi}(\mathcal{S}(w)) \neq \emptyset$, which is in contradiction with our supposition.

(b) $[\varphi] \cap min_{\subseteq}^{\varphi}(\mathcal{S}(w)) \neq \emptyset$. Recall that spheres in $\mathcal{S}(w)$ are totally ordered by set inclusion (nestedness condition) and that, by supposition, $[\psi] \cap min_{\subseteq}^{\psi}(\mathcal{S}(w)) \neq \emptyset$ and $min_{\subseteq}^{\varphi}(\mathcal{S}(w)) \neq min_{\subseteq}^{\psi}(\mathcal{S}(w))$. So, we have two cases to consider:

i. $min_{\subseteq}^{\varphi}(\mathcal{S}(w)) \supset min_{\subseteq}^{\psi}(\mathcal{S}(w))$; by assumption, since $[\psi] \cap min_{\subseteq}^{\psi}(\mathcal{S}(w)) \subseteq [\varphi]$ and $[\psi] \cap min_{\subseteq}^{\psi}(\mathcal{S}(w)) \neq \emptyset$, and since $v \in [\psi] \cap min_{\subseteq}^{\psi}(\mathcal{S}(w))$, it must be the case that the minimal φ-permitting sphere is contained in $min_{\subseteq}^{\psi}(\mathcal{S}(w))$, leading to a contradiction.

ii. $min_{\subseteq}^{\varphi}(\mathcal{S}(w)) \subset min_{\subseteq}^{\psi}(\mathcal{S}(w))$; by assumption, since $[\varphi] \cap min_{\subseteq}^{\varphi}(\mathcal{S}(w)) \subseteq [\psi]$ and $[\varphi] \cap min_{\subseteq}^{\varphi}(\mathcal{S}(w)) \neq \emptyset$, it must be the case that the minimal ψ-permitting sphere is contained in $min_{\subseteq}^{\varphi}(\mathcal{S}(w))$, leading to a contradiction.

($\supseteq$) We can reason similarly to the other direction ($\subseteq$).

So, $[\psi] \cap min_{\subseteq}^{\psi}(\mathcal{S}(w)) = [\varphi] \cap min_{\subseteq}^{\varphi}(\mathcal{S}(w))$. Therefore, by (i) and by definition of f, we have that $[\varphi]^f \cap min_{\subseteq}^{\varphi}(\mathcal{S}(w)) = [\psi]^f \cap min_{\subseteq}^{\psi}(\mathcal{S}(w))$, and so $f(\varphi, w) = f(\psi, w)$.

3. for all $\varphi, \psi \in \mathcal{L}_{\Box\!\rightarrow}$, for all $w \in W$, either $f(\varphi \vee \psi, w) \subseteq [\varphi]^f$ or $f(\varphi \vee \psi, w) \subseteq [\psi]^f$ or $f(\varphi \vee \psi, w) = f(\varphi, w) \cup f(\psi, w)$. Consider $[\varphi \vee \psi]$; by semantic conditions, we have that $[\varphi \vee \psi] = [\varphi] \cup [\psi]$. Assume that $f(\varphi \vee \psi) \nsubseteq [\varphi]^f$ and $f(\varphi \vee \psi) \nsubseteq [\psi]^f$. By (i) and definition of f, this means that $([\varphi] \cup [\psi]) \cap min_{\subseteq}^{\varphi\vee\psi}(\mathcal{S}(w)) \nsubseteq [\varphi]$ and $([\varphi] \cup [\psi]) \cap min_{\subseteq}^{\varphi\vee\psi}(\mathcal{S}(w)) \nsubseteq [\psi]$. Namely, there are $v, u \in ([\varphi] \cup [\psi]) \cap min_{\subseteq}^{\varphi\vee\psi}(\mathcal{S}(w))$ such that $v \notin [\varphi]$, i.e. $v \vDash \neg\varphi \wedge \psi$, and $u \notin [\psi]$, i.e. $u \vDash \varphi \wedge \neg\psi$. This also implies that, by the limit assumption, $min_{\subseteq}^{\varphi}(\mathcal{S}(w)) \neq \emptyset$ and $min_{\subseteq}^{\psi}(\mathcal{S}(w)) \neq \emptyset$.

Now, we are going to show that $[\varphi \vee \psi] \cap min_{\subseteq}^{\varphi\vee\psi}(\mathcal{S}(w)) = ([\varphi] \cap min_{\subseteq}^{\varphi}(\mathcal{S}(w))) \cup ([\psi] \cap min_{\subseteq}^{\psi}(\mathcal{S}(w)))$.

First, observe that $min_{\subseteq}^{\varphi}(\mathcal{S}(w)) \subseteq min_{\subseteq}^{\varphi\vee\psi}(\mathcal{S}(w))$. Indeed, suppose the contrary, then, by the nestedness condition, il holds that $min_{\subseteq}^{\varphi}(\mathcal{S}(w)) \supset min_{\subseteq}^{\varphi\vee\psi}(\mathcal{S}(w))$. However, by assumption,

$u \in min_{\subseteq}^{\varphi \vee \psi}(\mathcal{S}(w))$ and $u \vDash \varphi$. This means that the minimal φ-permitting cannot properly contain $min_{\subseteq}^{\varphi \vee \psi}(\mathcal{S}(w))$ contradicting the fact that $min_{\subseteq}^{\varphi}(\mathcal{S}(w)) \supset min_{\subseteq}^{\varphi \vee \psi}(\mathcal{S}(w))$.

Similarly, it holds that $min_{\subseteq}^{\varphi}(\mathcal{S}(w)) \supseteq min_{\subseteq}^{\varphi \vee \psi}(\mathcal{S}(w))$. Indeed, suppose the contrary, then, by the nestedness condition, il holds that $min_{\subseteq}^{\varphi}(\mathcal{S}(w)) \subset min_{\subseteq}^{\varphi \vee \psi}(\mathcal{S}(w))$. However, by assumption $u \in min_{\subseteq}^{\varphi \vee \psi}(\mathcal{S}(w))$ and $u \vDash \varphi$ and so $min_{\subseteq}^{\varphi}(\mathcal{S}(w))$ cannot be empty. This means that the minimal $\varphi \vee \psi$-permitting sphere must be contained in $min_{\subseteq}^{\varphi}(\mathcal{S}(w))$, contradicting the fact that $min_{\subseteq}^{\varphi}(\mathcal{S}(w)) \subset min_{\subseteq}^{\varphi \vee \psi}(\mathcal{S}(w))$.

Hence, it must holds that $min_{\subseteq}^{\varphi}(\mathcal{S}(w)) = min_{\subseteq}^{\varphi \vee \psi}(\mathcal{S}(w))$. Analogously we can reason with $min_{\subseteq}^{\psi}(\mathcal{S}(w))$ and v to show that $min_{\subseteq}^{\psi}(\mathcal{S}(w)) = min_{\subseteq}^{\varphi \vee \psi}(\mathcal{S}(w))$.

Therefore, since $[\varphi \vee \psi] = [\varphi] \cup [\psi]$, by the properties of set-theoretic operations, it must be the case that $[\varphi \vee \psi] \cap min_{\subseteq}^{\varphi \vee \psi}(\mathcal{S}(w)) = ([\varphi] \cap min_{\subseteq}^{\varphi}(\mathcal{S}(w))) \cup ([\psi] \cap min_{\subseteq}^{\psi}(\mathcal{S}(w)))$

Thus, by definition of f and Lemma 0.1, $f(\varphi \vee \psi, w) = f(\varphi, w) \cup f(\psi, w)$

□

Appendix B

Appendix of Chapter 2

B.1 Proof of Lemma 2.3

Proof. For a total spherical Lewisian model $\mathcal{M} = \langle W, \mathcal{S}, \vDash \rangle$, consider its associated restricted spherical model $\mathcal{M}_{\approx} = \langle W_{\approx}, \mathcal{S}_{\approx}, \vDash_{\approx} \rangle$.

1. $\mathcal{M}_{\approx}$ is a spherical total model.To show that $\mathcal{M}_{\approx}$ is a spherical Lewisian model, we must show that each system of spheres is nested. For $[w]_{\approx} \in I_{\approx}$ consider $S_{/\approx}, S'_{/\approx} \in \mathcal{S}_{\approx}([w]_{\approx})$. Since $\mathcal{M}$ is a spherical Lewisian model, we have that either $S \subseteq S'$ or $S' \subseteq S$. If the former is the case, then, by definition of $\mathcal{S}_{\approx}$, we have that $S_{/\approx} \subseteq S'_{/\approx}$. Analogously, if the latter is the case, $S_{/\approx} \supseteq S'_{/\approx}$.

   Now, consider a formula $\varphi \in \mathcal{L}$ such that $\vdash_{CPL} \neg\varphi$ and take any $[w]_{\approx} \in W_{\approx}$. We know that $\mathcal{S}_{\approx}([w]_{\approx}) = \{S_{/\equiv} : S \in \mathcal{S}(c([w]_{\approx}))\}$. By the totality of $\mathcal{M}$, we have that $\bigcup \mathcal{S}(c([w]_{\approx})) \cap [\varphi] \neq \emptyset$, and so we immediately get that $\mathcal{S}_{\approx}([w]_{\approx}) \cap [\varphi]_{\approx} \neq \emptyset$. Hence

2. if $\mathcal{M}$ satisfies **Centering** then $\mathcal{M}_{\approx}$ satisfies **Centering** too. Immediate from the definition of $\mathcal{S}_{\approx}$.

3. Consider the function $E : Val_{CPL} \rightarrowtail\!\!\!\!\twoheadrightarrow W_{\approx}$ defined as:

   $$E(v) = \{w \in W \mid w \vDash p \Leftrightarrow v(p) = 1 \text{ and } w \nvDash p \Leftrightarrow v(p) = 0\}$$

   We re going to show that E is:

   - *Well defined*. The codomain of E is indeed $W_{\approx}$. In fact, consider any $w \in E[v]$. We know that there must be such w, i.e. $E(v) \neq \emptyset$ by totality and the fact that Var is finite. Indeed, consider the formula $\bigwedge_{p:v(p)=1} p \wedge \bigwedge_{\neg p:v(p)=0} \neg p$. This formula is satisfiable and so, by totality, there must be a world in W at which this formulas is true. Furthermore, consider any $v \in W$ such that $w \approx v$. Then, w and v are such that for all $p \in Var$, $w \models p \Leftrightarrow v \models p$. Therefore for all $p \in Var$, $v(p) = 1 \Leftrightarrow v \models p$, hence $v \in E(v)$. Namely $E(v) = [w]_{\approx}$.

- *Injective*. Consider $v_1, v_2 \in Val$ such that $v_1 \neq v_2$. Thus, there must be a $p \in Var$ such that, without loss of generality, $v_1(p) = 1$ and $v_2(p) = 0$. Now, consider a $w \in W$ such that for all $p \in Var$, $w \models p \Leftrightarrow v_1(p) = 1$. By totality, we know that such w exists. It is the case that $w \in E(v_1)$, but, clearly, $w \notin E(v_2)$. So $E(v_1) \neq E(v_2)$
- *Suijective*. Consider any $[w]_{\approx}$. First, notice that for any other $[v]_{\approx} \in W_{\approx}$, there must be a $p \in Var$, such that, without loss of generality, $[w]_{\approx} \models p$ but $[v]_{\approx} \not\models p$, otherwise $[w]_{\approx} = [v]_{\approx}$. Hence, $[\bigwedge_{p:w\models_{\approx}} p \wedge \bigwedge_{\neg p:w\not\models_{\approx} p} \neg p]_{\approx}$ is a singleton. Now, consider the valuation v defined as: for all $p \in Var$, $v(p) = 1 \Leftrightarrow [w]_{\approx} \models_{\approx} p$. It follows that $E(v) = [w]_{\approx}$.

Therefore, there is a bijection between Val_{CPL} and $W_{\approx}$, and so $W_{\approx} = 2^n$ where $n = |Var|$. Moreover, notice that, by totality, for all $v \in Val$, for all $[w]_{\approx} \in W_{\approx}$,

$$[\bigwedge_{p:w\models_{\approx}} p \wedge \bigwedge_{\neg p:w\not\models_{\approx} p} \neg p]_{\approx} \cap \bigcup \mathcal{S}_{\approx}([w]_{\approx}) \neq \emptyset$$

and since $[\bigwedge_{p:w\models_{\approx}} p \wedge \bigwedge_{\neg p:w\not\models_{\approx} p} \neg p]_{\approx}$, we have that $|\bigcup \mathcal{S}_{\approx}([w]_{\approx})| = |W_{\approx}| = 2^n = |Val_{CPL}|$.

□

B.2 Proof of Lemma 2.18

Proof.

1. Trivial by construction of the partition Π.
2. If $c_l \neq r$ it might be the case that there exists $\beta \in c_l$ and $\beta \notin r$. However, if this is the case, then, by cardinality argument and the construction of $\mathbf{C}_l$, there must be $\beta' \in r$ such that $\beta' \notin c_l$.
3. For reduction, assume that there is $\alpha_{i,j} \in r$ such that $\alpha_{i,j} \notin c_l$ and for all $\beta \in c_l$, β appears in the j-th row in some preceding column with respect to the i-th column. Now, since $\mathbf{C}_l$ is the first submatrix of $\mathbf{R}[\omega]$ for which the spherical property fails, all these β's cannot appear before $\mathbf{c}_l$ in the indexing of $\mathbf{R}[\omega]$, for otherwise there would exists a $\mathbf{C}_h$ with $h < l$ where the relevant spherical property would fail, contradicting the fact that $\mathbf{C}_l$ is the first submatrix of $\mathbf{R}[\omega]$ for which the spherical property fails.

Thus, these β's must necessarily belong to $\mathbf{r}$ and appear earlier than $\alpha_{i,j}$. This implies that $\mathbf{r}$ has length at least $|c_l| + 1$ as it contains all the elements of c_l plus $\alpha_{i,j}$ that does not belong to $\mathbf{c}_l$ by assumption. This is in contradiction with the construction of $\mathbb{C}_\omega$ and hence of $\mathbf{C}_l$.

□

B.3 Proof of Lemma 2.20

Proof. Consider the $LBC_\Box$-valuation, $\langle Perm(Val_{CPL}), R, \vDash\rangle$, induced by a restricted canonical model $\mathcal{M}^{\mathfrak{C}} = \langle Val_{CPL}, \mathcal{S}^{\mathfrak{C}}, \vDash^{\mathfrak{C}}\rangle$. The R is such that:

- R satisfies **Ser** in Table 2.3.2. This property clearly holds since $\mathcal{M}^{\mathfrak{C}}$ is total, hence for all $v \in Val_{CPL}$, $\bigcup \mathcal{S}^{\mathfrak{C}}(v) \neq \emptyset$. This means that $Radius_v \neq \emptyset$ (see Definition 2.22). Hence, for all $e \in Perm(Val_{CPL})$, $R[e] \neq \emptyset$

- R satisfies **Cen** in Table 2.3.2. This properties follows from the fact that $\mathcal{M}^{\mathfrak{C}}$ satisfies **Centering** in Table 1. Hence, for all $v \in Val_{CPL}$, $\{v\} \in \mathcal{S}^{\mathfrak{C}}(v)$. This means that for all $r, r' \in Radius_v$, $r[1] = r'[1]$. Hence, since for all e, $R[e]$ is (up to bijective correspondence) $Radius_x$, for some $x \in Val_{CPL}$, then R satisfies **Cen**, i.e. for all $e \in Perm(Val_{CPL})$, for all $e^1, e^2 \in R[e]$, $e^1[1] = e^2[1]$.

- R satisfies **Sph** in Table 2.3.2. Specifically, we will show that for all $e \in Perm(Val_{CPL})$, $\mathbf{R}[e]$ is spherical. Observe that (up to isomorphism) $\mathbf{R}[e]$ is made of the elements of $Radius_x$, for some $x \in Val_{CPL}$. Specifically, the rows of $\mathbf{R}[e]$ can be identified with elements of $Radius_x$. Now, observe the definition of $Radius_x$ from Definition 2.22. Every elements $r \in Radius_x$, is such that

  $$r \in Perm(S^x_1) \times Perm(S^x_2 \setminus S^x_1) \times Perm(S^x_3 \setminus S^x_2) \times \cdots \times Perm(S^x_n \setminus S^x_{n-1})$$

  Now, let $\mathbf{Perm}(X)$ denote the matrix whose rows are exactly the elements of $Perm(X)$. Consider the partition Π of $\mathbf{R}[e]$ such that $\Pi = \{\mathbf{Perm}(S^x_1), \mathbf{Perm}(S^x_2 \setminus S^x_1), \ldots, \mathbf{Perm}(S^x_n \setminus S^x_{n-1})\}$. It is clear that, by construction of $\mathbf{Perm}(X)$ and definition of permutation, for all $\mathbf{Perm}(X) \in \Pi$, all the elements appearing in the first column of $\mathbf{Perm}(X)$ are exactly the same as the elements appearing in each of the rows of $\mathbf{Perm}(X)$. Hence, $\langle Perm(Val_{CPL}), R, \rangle$ is **Spherical**.

□

B.4 Proof of Lemma 2.22

Proof. To prove this, consider $x \in Val_{CPL}$ and assume $\mathcal{S}^{\mathfrak{C}} = \{S_1, \dots, S_t\}$, where $S_1 = \{x\}$ and $S_1 \subseteq S_2 \subseteq \cdots \subseteq S_t$. By definition of $f^{\mathfrak{C}}$ and by totality, there will be a least index l such that $f^{\mathfrak{C}}(\varphi, x) \subseteq S_l$. More formally, by definition of $f^{\mathfrak{C}}$ and by totality of $\mathcal{M}^{\mathfrak{C}}$, there is a $S_l \in \mathcal{S}^{\mathfrak{C}}(x)$ such that $min_{\varphi}(\mathcal{S}^{\mathfrak{C}}(x)) = S_l$. By the construction put forward in Definition 2.22 that leads from spherical models to Lewis frames, we can observe that there is a spherical partition $\Pi = \{\mathbf{C}_1, \dots, \mathbf{C}_t\}$ of $\mathbf{R}[d(x)]$ such that for all $1 < m \leq t$, $C_m = S_m \setminus S_{m-1}$ and $C_1 = S_1$, since elements of S_i are classical valuations. Hence, there is C_l such that $\{v \in Val_{CPL} \mid v(\varphi) = 1\} \cap C_l \neq \emptyset$ and for all $k < l$, $C_k \cap \{v \in Val_{CPL} \mid v(\varphi) = 1\} = \emptyset$. Namely, C_l is the least set containing a classical valuation making φ true, and moreover $C_l = S_l \setminus S_{l-1}$ (or $C_l = S_1$ if $l = 1$). Moreover, by **Sphericality**, we also know that $C_l = c_l^1$ where $\mathbf{c}_l^1$ is the first column of the submatrix $\mathbf{C}_l$. Notice, also, that rows of $\mathbf{C}_l$ are permutations of elements of $\mathbf{c}_l^1$, hence we have that:

$$f([\varphi]_{\dashv\vdash_{CPL}}, d(x)) = C_l \cap \{v \in Val_{CPl} \mid v(\varphi) = 1\} = S_l \cap [\varphi]^{\mathfrak{C}}$$

This equality holds since $C_l = S_l \setminus S_{l-1}$ (or $C_l = S_1$ if $l = 1$), $[\varphi]^{\mathfrak{C}} = \{v \in Val \mid v(\varphi) = 1\}$, and since we know that $C_l = c_l^1 \cap \{v \in Val_{CPL} \mid v(\varphi) =\} \neq \emptyset$ and that rows of $\mathbf{C}_l$ are permutations of elements of $\mathbf{c}_l^1$.

As a corollary of the above equality, we get that $f^{\mathfrak{C}}(\varphi, x) = f([\varphi]_{\dashv\vdash_{CPL}}, d(x))$ □

Appendix C

Appendix of Chapter 3

C.1 Proof of Lemma 3.8

Proof. The idea is that by construction of $Radius_w$ in Definition 3.8, and by the structure of r, $g(\varphi, r)$ would always coincide with $Sel(\varphi, r)$ except possibly when $Sel(\varphi, r) = r[1]$. We unpack the proof into three cases:

1. $g(\varphi, r) = \{r\}$; by Lemma 3.7, we have that $r \equiv r[1][1]$, and so, clearly, $Sel(\varphi, r) = \{r[1]\}$, namely $r[1]$ is the first φ-world appearing in r. Hence, since $r \equiv r[1]$, it is the case that $g(\varphi, r)_{/\equiv} = Sel(\varphi, r)_{/\equiv}$. We can reason similarly if $Sel(\varphi, r) = \{r[1]\}$ and show that $g(\varphi, r)_{/\equiv} = Sel(\varphi, r)_{/\equiv}$.

2. $g(\varphi, r) = \emptyset$; this means that $min^{\varphi}_{\subseteq}(\mathcal{S}^{\mathfrak{R}}(r)) = \emptyset$; hence, by case 1 above and by construction of radiation, there cannot be any i such that $r[i] \in [\varphi]^{\mathfrak{R}}$. Hence $Sel(\varphi, r) = \emptyset = g(\varphi, r)$. We can reason analogously if $Sel(\varphi, r) = \emptyset$ and show that $g(\varphi, r) = Sel(\varphi, r)$

3. $g(\varphi, r) \neq \emptyset$ and $g(\varphi, r) \neq \{r\}$. In this cases, we show that $g(\varphi, w) = Sel(\varphi, r)$. Specifically, by assumption, there is a $S^r_i \in \mathcal{S}^{\mathfrak{R}}(r)$ such that $S^r_i, = min^{\varphi}_{\subseteq}(\mathcal{S}^{\mathfrak{R}}(r))$. Observe that by construction it must be the case that $[\varphi]^{\mathfrak{R}} \cap min^{\varphi}_{\subseteq}(\mathcal{S}^{\mathfrak{R}}(r)) = \{r[i]\}$, since $S^r_i \setminus (\bigcup_{S^r_k \subset S^r_i} S^r_k) = \{r[i]\}$ and $S^r_i = min^{\varphi}_{\subseteq}(\mathcal{S}^{\mathfrak{R}}(r))$. In words, $r[i]$ is the only element that belongs to S^r_i but doesn't belong to all the other spheres smaller than S^r_i; hence, since S^r_i is the minimal φ-permitting sphere in $\mathcal{S}^{\mathfrak{R}}(r)$, we have that $r[i] \vDash^{\mathfrak{R}} \varphi$, thus $[\varphi]^{\mathfrak{R}} \cap min^{\varphi}_{\subseteq}(\mathcal{S}^{\mathfrak{R}}(r)) = \{r[i]\}$. Therefore, clearly $r[i] \in [\varphi]^{\mathfrak{R}}$ and that for all $1 \leq k < i$, $r[k] \notin$. Indeed, if there were such k, it would be the case that $S^r_k \cap [\varphi]^{\mathfrak{R}} \neq \emptyset$ and $S^r_k \subseteq S^r_i$, contradicting the assumption that S^r_i is the minimal φ-permitting sphere in $\mathcal{S}^{\mathfrak{R}}(r)$. Hence, we have that $g(\varphi, r) = Sel(\varphi, r)$. Conversely, assume that $Sel(\varphi, r) \neq \emptyset$ and $Sel(\varphi, r) \neq \{r[1]\}$, namely there is $i > 1$ such that $r[i] \in [\varphi]^{\mathfrak{R}}$. Consider S^r_i: it is

straightforward to see that $S_i^r = min_{\subseteq}^{\varphi}(\mathcal{S}^{\mathfrak{R}}(r))$ and that $[\varphi]^{\mathfrak{R}} \cap S_i^r = \{r[i]\}$, and so $g(\varphi, r) = Sel(\varphi, r)$.

Hence, it must hold that $g(\varphi, r)_{/\equiv} = Sel(\varphi, r)_{/\equiv}$.

□

C.2 Proof of Theorem 3.5

Proof. The soundness proof is straightforward and proceed as customary. For the completeness proof we will use a canonical model construction. A canonical model for a **KV**$\mathfrak{C}$ logic is a tuple

$$\langle MCS, R, g, f, \vDash \rangle$$

where

- **MCS** is the set of all maximally consistent sets of sentences in $\mathcal{L}_{>}^{\Box}$, such that, for all $x \in$ **MCS**, if $x \vDash_{\mathbf{KV}\mathfrak{C}} \varphi$, then $\varphi \in x$
- R is defined as in the canonical model for the completeness proof of normal modal logics: for $x, y \in MCS$,

$$xRy \Leftrightarrow \text{for all } \varphi \in For_{\mathcal{L}_{>}^{\Box}}, \text{ if } \Box\varphi \in x, \text{ then } \varphi \in y$$

- $g : For_{\mathcal{L}_{>}^{\Box}} \times MCS \rightarrow \wp(MCS)$ is defined as follows:

  $g(\varphi,)$ is the set of all maximally consistent set that extend $\{\psi \mid \varphi > \psi\}$

- $f : For_{\mathcal{L}_{!>}^{\Box}} \times MCS \rightarrow \wp(MCS)$ is defined as follows:

$$f(\varphi, x) = \bigcup_{y \in R[x]} g(\varphi, y)$$

- $\vDash = \in$, i.e. $x \vDash \varphi \Leftrightarrow \varphi \in x$

It is not difficult to show that the above model is indeed a spherical Kripke model and $\vDash$ behaves exactly as in Definition 3.7, or more precisely, that the spherical Lewisian model induced by the above model is a spherical Kripke model. The idea is that we can work with a function g instead of a Stalnakerian system of spheres. This indeed would simplify the proof. Then, using the procedure in (Lewis 1971) or (Lewis 1973b), we can convert the above model into a one in which g is replaced by an equivalent Stalnakerian system of sphere $\mathcal{S}$:

$$\langle MCS, R, \mathcal{S}, f, \vDash \rangle$$

and R remains unchanged. Then, it is easy to show that $g(\varphi x) = [\varphi] \cap min_{\varphi}(\mathcal{S}(x))$, and so, since R is also unchanged, the two models will be equivalent in that elements of MCS, in the new converted model, will force exactly the same formulas as in the original model. The fact that $\models$ behaves as desired follows from the property of the canonical models for modal logic (Blackburn, de Rijke, and Venema 2001) and the property of the canonical models for variably strict conditional logics in (Lewis 1971, 1973b). The important step consist in showing that f forces the **Sphericality** condition in Definition 3.7. The idea is the following: for $i \in I$, we know that axiom $(KV1)$ is in i. By maximal consistency, we know that at least one of the disjuncts of axiom $(KV1)$ is in i. Suppose, $\Box((\varphi \lor \psi) > \varphi) \in i$. Then, by semantic conditions, $(\varphi \lor \psi) > \varphi \in j$ for all $j \in R[i]$. By semantic conditions, this implies that $g(\varphi \lor \psi, j) \subseteq [\psi]$ for all $j \in R[i]$, hence, by definition of f, $f(\varphi \lor \psi, i) \subseteq [\varphi]$. Analogously for the other conjunct. The only interesting conjunct is the more complex one containing the bi-conditional, namely $\Box((\varphi \lor \psi) > \varphi) \lor \Box((\varphi \lor \psi) > \psi) \lor (\Box((\varphi \lor \psi) > \delta)) \leftrightarrow \Box((\varphi > \delta) \land (\psi > \delta)))$ for all $\varphi, \psi, \delta \in For_{\mathcal{L}_{\restriction >}^{\Box}}$. The key observation is that this axiom implies that for all $\delta \in For_{\mathcal{L}_{\restriction >}^{\Box}}$, if $f(\varphi \lor \psi, i) \subseteq [\delta]$, then $f(\varphi, i) \subseteq [\delta]$ and $f(\psi, i) \subseteq [\delta]$. Indeed, assume $f(\varphi \lor \psi, i) \subseteq [\delta]$, by definition of f, this means that for all $j \in R[i]$, $g(\varphi \lor \psi) \subseteq [\delta]$. Hence, this means that for all $j \in R[i]$, $(\varphi \lor \psi) > \delta \in j$, and so $\Box((\varphi \lor \psi) > \delta) \in i$. By the fact that the relevant conjunct of axiom $(KV1)$ is in i, then we also have that $\Box(\varphi > \delta) \in i$ and $\Box(\psi > \delta) \in i$. And so $\varphi > \delta \in j$ and $\psi > \delta \in j$ for all $j \in R[i]$. Therefore, by semantic condition $g(\varphi, j) \subseteq [\delta]$ and $g(\psi, j) \subseteq [\delta]$ for all $j \in R[i]$. By definition of f, this means that $f(\varphi, i) \subseteq [\delta]$ and $f(\psi, i) \subseteq [\delta]$. Also the other direction holds analogously, since, if $f(\varphi) \cup f(\psi) \subseteq [\delta]$, then $\Box(\varphi > \delta) \in i$ and $\Box(\psi > \delta) \in i$, so, by axiom $(KV1)$, we also have that $\Box((\varphi \lor \psi) > \delta)$, and so $f(\varphi \lor \psi, i) \subseteq [\delta]$. Hence, we must have that $f(\varphi, i) \cup f(\psi, i) \subseteq [\delta]$ iff $f(\varphi \lor \psi, i) \subseteq [\delta]$ for all $\delta \in For_{\mathcal{L}_{\restriction >}^{\Box}}$, and all $\varphi, \psi \in For_{\mathcal{L}_{\restriction >}^{\Box}}$. Additionally, notice that $f(\varphi, i)$ is the set of maximally consistent sets that extends $\{\psi \mid \Box(\varphi > \psi) \in i\}$. By what we showed before, we have that for all counterfactual formulas δ, $\{\psi \mid \Box(\varphi \lor \psi > \delta) \in i\} \vdash_{\mathbf{KV}\mathfrak{C}} \delta$ iff $\{\psi \mid \Box(\varphi > \psi) \in i\} \cap \{\psi \mid \Box(\varphi > \psi) \in i\} \vdash_{\mathbf{KV}\mathfrak{C}} \delta$. So, basically, we have two sets that derive exactly the same formulas in the fragment $For_{\mathcal{L}_{\restriction >}^{\Box}}$. They could only eventually differ if the contain two different formulas of the $\varphi > \psi$, which are not in the fragment $For_{\mathcal{L}_{\restriction >}^{\Box}}$. Moreover, all their maximally consistent extensions, $f(\varphi \lor \psi, i)$ and $f(\varphi, i) \cup f(\psi, i)$, must agree on the same formulas in the fragment $For_{\mathcal{L}_{\restriction >}^{\Box}}$. Hence, $f(\varphi \lor \psi, i)_{/\equiv} = f(\psi, i)_{/\equiv} \cup f(\varphi, i)_{/\equiv}$. Namely f satisfies **Sphericality**. Moreover, by Lemma 3.3, we have that the canonical model for a logic $\mathbf{KV}\mathfrak{C}$ satisfies condition $\mathfrak{C}$.

Now, by contraposition, assume $\Gamma \nvdash_{\mathbf{KV}\mathfrak{C}} \varphi$. Then Γ can be extended to a maximally consistent set x in MCS such that $x \vDash \gamma$ for all $\gamma \in \Gamma$ and $x \nvDash \varphi$. Moreover, by the results in (Lewis 1971, pp.78-79) and in (Lewis 1973b, pp. 58-59), our canonical model can be converted into a model of the form $\langle MCS, R, \mathcal{S}, f, \vDash \rangle$ in which g is induced by $\mathcal{S}$, and so f also remain the same. Hence, since $x \vDash \gamma$ for all $\gamma \in \Gamma$ and $x \nvDash \varphi$, we also have that $\Gamma \not\models_{\mathbf{KV}\mathfrak{C}} \varphi$

□

Bibliography

Adams, Ernest W. (1965), "The Logic of Conditionals", *Inquiry*, 8, pp. 166-197.

— (1966), "Probability and the Logic of Conditionals", in *Aspects of Inductive Logic*, Elsevier, pp. 265-316.

— (1970), "Subjunctive and Indicative Conditionals", *Foundations of Language*, 6, 1, pp. 89-94.

— (1975), *The Logic of Conditionals*, Reidel, Dordrecht.

— (1976), "Prior Probabilities and Counterfactual Conditionals", in *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*, Springer Netherlands, pp. 1-21.

— (1993), "On the rightness of certain counterfactuals", *Pacific Philosophical Quarterly*, 74, 1 (Mar. 1993), pp. 1-10.

Barbero, Fausto and Gabriel Sandu (2020), "Team Semantics for Interventionist Counterfactuals: Observations vs. Interventions", *Journal of Philosophical Logic*, 50, 3 (Dec. 2020), pp. 471-521.

Bennett, Jonathan (2003), *A Philosophical Guide to Conditionals*, Oxford University PressOxford.

Berto, Francesco, Rohan French, Graham Priest, and David Ripley (2017), "Williamson on Counterpossibles", *Journal of Philosophical Logic*, 47, 4 (Aug. 2017), pp. 693-713.

Blackburn, Patrick, Maarten de Rijke, and Yde Venema (2001), *Modal Logic*, Cambridge University Press.

Blok, W. J. and Don Pigozzi (2014), *Algebraizable Logics*, Advanced Reasoning Forum.

Briggs, R.A. (2012), "Interventionist Counterfactuals", *Philosophical Studies*, 160, 1, pp. 139-166.

Burris, S. and H.P. Sankappanavar (1981), *A Course in Universal Algebras*, Springer-Verlag.

Byrne, Ruth M. J. (2019), "Counterfactuals in Explainable Artificial Intelligence (XAI): Evidence from Human Reasoning", in *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence*, International Joint Conferences on Artificial Intelligence Organization.

Canavotto, Ilaria and Eric Pacuit (2021), "Choice-Driven Counterfactuals", *Journal of Philosophical Logic*, 51, 2 (Oct. 2021), pp. 297-345.

Carrara, Massimiliano, Davide Fazio, and Michele Pra Baldi (2022), "Paraconsistent Belief Revision: An Algebraic Investigation", *Erkenntnis* (Apr. 2022).

Chellas, Brian F. (1975), "Basic conditional logic", *Journal of Philosophical Logic*, 4, 2 (May 1975), pp. 133-153.

Chou, Yu-Liang, Catarina Moreira, Peter Bruza, Chun Ouyang, and Joaquim Jorge (2022), "Counterfactuals and causability in explainable artificial intelligence: Theory, algorithms, and applications", *Information Fusion*, 81 (May 2022), pp. 59-83.

Ciardelli, Ivano (2021), "Restriction without Quantification: Embedding and Probability for Indicative Conditionals", *Ergo an Open Access Journal of Philosophy*, 8, 0 (Dec. 2021).

— (2022), *Inquisitive Logic*, Springer International Publishing.

Ciardelli, Ivano, Linmin Zhang, and Lucas Champollion (2018), "Two switches in the theory of counterfactuals", *Linguistics and Philosophy*, 41, 6 (June 2018), pp. 577-621.

Coletti, Giulianella and Marcello Mastroleo (2006), "Conditional Belief Functions: a Comparison among Different Definitions" (Jan. 2006).

Crupi, Vincenzo and Andrea Iacona (2020), "The Evidential Conditional", *Erkenntnis*, 87, 6 (Dec. 2020), pp. 2897-2921.

Dalen, Dirk Van (1986), "Intuitionistic Logic", in *Handbook of Philosophical Logic*, Springer Netherlands, pp. 225-339.

Davey, B. A. and H. A. Priestley (2002), *Introduction to Lattices and Order*, Cambridge University Press.

Dempster, Arthur (1968), "A Generalisation of Bayesian Inference", *Journal of the Royal Statistical Society Series B*, 30, pp. 205-247.

Denoeux, Thierry (2011), "Introduction to Belief Functions", presentation at he 1st school on belief functions and their applications, Autrans.

Díez, José (2015), "Counterfactuals, the Discrimination Problem and the Limit Assumption", *International Journal of Philosophical Studies*, 23, 1 (Jan. 2015), pp. 85-110.

Douven, Igor and Sara Verbrugge (2010), "The Adams family", *Cognition*, 117, 3, pp. 302-318.

— (2013), "The Probabilities of Conditionals Revisited", *Cognitive Science*, 37, 4, pp. 711-730.

Dubois, Didier and Henri Prade (1994), "A survey of belief revision and updating rules in various uncertainty models", *International Journal of Intelligent Systems*, 9, 1, pp. 61-100.

Edgington, Dorothy (1995), "On Conditionals", *Mind*, 104, 414, pp. 235-329.

— (2008), "I-Counterfactuals", *Proceedings of the Aristotelian Society (Hardback)*, 108, 1part1 (Mar. 2008), pp. 1-21.

Égré, Paul and Hans Rott (2021), "The Logic of Conditionals", in *The Stanford Encyclopedia of Philosophy*, ed. by Ed Zalta, Retrieved from https://plato.stanford.edu/archives/fall2021/entries/logic-conditionals/, Stanford Metaphysics Lab.

Esakia, Leo (2019), *Heyting Algebras*, ed. by Guram Bezhanishvili and Wesley H. Holliday, Springer International Publishing.

Eva, Benjamin, Reuben Stern, and Stephan Hartmann (2019), "The Similarity of Causal Structure", *Philosophy of Science*, 86, 5, pp. 821-835.

Evans, Jonathan, Simon Handley, Helen Neilens, and David Over (2007), "Thinking about conditionals: A study of individual differences", *Memory & Cognition*, 35, 7 (Oct. 2007), pp. 1772-1784.

Fine, Kit (1975), "Critical Notice of Lewis, Counterfactuals", *Mind*, 84, 335, pp. 451-458.

— (2012), "Counterfactuals Without Possible Worlds", *Journal of Philosophy*, 109, 3, pp. 221-246.

— (2016), "Angellic Content", *Journal of Philosophical Logic*, 45, 2, pp. 199-226.

— (2017), "Truthmaker Semantics", in *A Companion to the Philosophy of Language*, ed. by Bob Hale, Crispin Wright, and Alexander Miller, Wiley, New York, pp. 556-577.

Flaminio, Tommaso, Lluis Godo, and Hykel Hosni (2020), "Boolean algebras of conditionals, probability and logic", *Artificial Intelligence*, 286 (Sept. 2020), p. 103347.

Font, Josep (2016), *Abstract Algebraic Logic - An Introductory Textbook*.

Font, Josep M. and Gonzalo Rodríguez (1990), "Note on algebraic models for relevance logic", *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 36, 6, pp. 535-540.

Galles, David and Judea Pearl (1998), "An Axiomatic Characterization of Causal Counterfactuals", *Foundations of Science*, 3, pp. 151-182.

Gärdenfors, Peter (1982), "Imaging and Conditionalization", *Journal of Philosophy*, 79, pp. 747-760.

Gibbard, Allan (1980), "Two Recent Theories of Conditionals", in *IFS*, Springer Netherlands, pp. 211-247.

Ginsberg, Matthew L. (1986), "Counterfactuals", *Artificial Intelligence*, 30, 1 (Oct. 1986), pp. 35-79.

Girlando, Marianna, Sara Negri, and Nicola Olivetti (2021), "Uniform labelled calculi for preferential conditional logics based on neighbourhood semantics", *Journal of Logic and Computation*, 31, 3 (Apr. 2021), pp. 947-997.

Gödel, Kurt (1986), *Collected Works, Volume 1: Publications 1929-1936*, Oxford, England: Clarendon Press.

Günther, Mario (2017), "Disjunctive Antecedents for Causal Models", in *Proceedings of the 21st Amsterdam Colloquium*.

— (2018), "Learning Conditional Information by Jeffrey Imaging on Stalnaker Conditionals", *Journal of Philosophical Logic*, 47, pp. 851-876.

Günther, Mario (2022), "Probabilities of Conditionals and Imaged Probabilities", Draft manuscript.

Hájek, Alan (1989), "Probabilities of conditionals - Revisited", *Journal of Philosophical Logic*, 18, 4 (Nov. 1989).

— (2014), "Probabilities of counterfactuals and counterfactual probabilities", *Journal of Applied Logic*, 12, 3 (Sept. 2014), pp. 235-251.

Hájek, Petr (1996), "Getting Belief Functions from Kripke Models", *International Journal of General Systems*, 24, 3 (Feb. 1996), pp. 325-327.

Hale, Bob and Aviv Hoffmann (eds.) (2010), *Modality*, Oxford University PressOxford.

Halmos, Paul and Steven Givant (2009), *Introduction to Boolean Algebras*, Springer New York.

Halpern, Joseph Y. (2000), "Axiomatizing Causal Reasoning", *Journal of Artificial Intelligence Research*, 12, pp. 317-337.

— (2013), "From Causal Models to Counterfactual Structures", *Review of Symbolic Logic*, 6, 2, pp. 305-322.

— (2016), *Actual Causation*, MIT Press.

Halpern, Joseph Y. and Judea Pearl (2005), "Causes and Explanations: A Structural-Model Approach. Part I: Causes", *British Journal for the Philosophy of Science*, 56, pp. 843-887.

Harmanec, David, George J. Klir, and Germano Resconi (1994), "On modal logic interpretation of Dempster-Shafer theory of evidence", *International Journal of Intelligent Systems*, 9, 10, pp. 941-951.

Harmanec, David, George J. Klir, and Zhenyuan Wang (1996), "Modal logic interpretation of Dempster-Shafer theory: An infinite case", *International Journal of Approximate Reasoning*, 14, 2-3 (Feb. 1996), pp. 81-93.

Hazen, A. (1982), "On a possible misinterpretation of Kripke's semantics for intuitionistic logic", *Analysis*, 42, 3 (June 1982), pp. 128-133.

Herzberger, Hans G. (1979), "Counterfactuals and Consistency", *The Journal of Philosophy*, 76, 2 (Feb. 1979), p. 83.

Heyting, A. (1930), "Die formalen Regeln der intuitionistischen Logik I, II, III", *Sitzungsberichte derpreussischen Akademie von Wissenschaften*, pp. 42-56, 57-71, 158-169.

Holliday, Wesley H., Tomohiro Hoshi, and Thomas F. Icard (2013), "Information dynamics and uniform substitution", *Synthese*, 190, S1 (Apr. 2013), pp. 31-55.

Iatridou, Sabine (2000), "The Grammatical Ingredients of Counterfactuality", *Linguistic Inquiry*, 31, 2 (Apr. 2000), pp. 231-270.

Ippolito, Michela (2006), "Semantic Composition and Presupposition Projection in Subjunctive Conditionals", *Linguistics and Philosophy*, 29, 6 (Dec. 2006), pp. 631-672.

— (2013), *Subjunctive Conditionals: A Linguistic Analysis*, Cambridge, MA: MIT Press.

Isard, S. DD (1974), "What Would You Have Done If...?", *Theoretical Linguistics*, 1, 1-3.

Jansana, Ramon (1995), "Abstract modal logics", *Studia Logica*, 55, 2, pp. 273-299.

Jipsen, Peter (1992), *Computer Aided Investigations of Relation Algebars*, PhD thesis, Vanderbilt University.

Jónsson, Bjarni and Alfred Tarski (1951), "Boolean Algebras with Operators. Part I", *American Journal of Mathematics*, 73, 4 (Oct. 1951), p. 891.

Joyce, James (1999), *Foundations of Causal Decision Theory*, Cambridge University Press, Cambridge.

Kaufmann, Stefan (2013), "Causal Premise Semantics", *Cognitive Science*, 37, 6, pp. 1136-1170.

— (2017), "The Limit Assumption", *Semantics and Pragmatics*, 10, 18 (Dec. 2017), p. 1.

Khoo, Justin (2015), "On Indicative and Subjunctive Conditionals", *Philosophers' Imprint*, 15.

Kracht, Marcus (1999), *Tools and Techniques in Modal Logic*, New York: Elsevier.

Kreisel, G. (1970), "Saul A. Kripke. Semantical analysis of intuitionistic logic I. Formal systems and recursive functions, Proceedings of the Eighth Logic Colloquium, Oxford, July 1963, edited by J. N. Crossley and M. A. E. Dummett, Series in logic and the foundations of mathematics, North-Holland Publishing Company, Amsterdam1965, pp. 92–130." *Journal of Symbolic Logic*, 35, 2 (June 1970), pp. 330-332.

Kripke, Saul A. (1965), "Semantical Analysis of Intuitionistic Logic I", in *Formal Systems and Recursive Functions*, Elsevier, pp. 92-130.

Leitgeb, Hannes (2011), "A probabilistic semantics for counterfactuals, Part A", *The Review of Symbolic Logic*, 5, 1 (Nov. 2011), pp. 26-84.

Lewis, C. I. (1912), "Implication and the Algebra of Logic", *Mind*, XXI, 84, pp. 522-531.

Lewis, David (1971), "Completeness and decidability of three logics of counterfactual conditionals1", *Theoria*, 37, 1 (Feb. 1971), pp. 74-85.

— (1973a), "Causation", *Journal of Philosophy*, 70, 17, pp. 556-567.

— (1973b), *Counterfactuals*, Blackwell, Oxford.

— (1973c), "Counterfactuals and Comparative Possibility", *Journal of Philosophical Logic*, 2, 4, pp. 418-446.

— (1976), "Probabilities of Conditionals and Conditional Probabilities", *Philosophical Review*, 85, 3, pp. 297-315.

— (1979), "Counterfactual Dependence and Time's Arrow", *Noûs*, 13, pp. 455-476.

Loewer, Barry (1976), "Counterfactuals with Disjunctive Antecedents", *Journal of Philosophy*, 73, 16, pp. 531-537.

Lyons, John (1977), *Semantics*, Cambridge University Press.

McKay, Thomas and Peter Van Inwagen (1977), "Counterfactuals with Disjunctive Antecedents", *Philosophical Studies*, 31, pp. 353-356.

McKinsey, J. C. C. and Alfred Tarski (1948), "Some theorems about the sentential calculi of Lewis and Heyting", *Journal of Symbolic Logic*, 13, 1 (Mar. 1948), pp. 1-15.

McLoone, Brian (2020), "Calculus and counterpossibles in science", *Synthese*, 198, 12 (Sept. 2020), pp. 12153-12174.

Miller, Tim (2019), "Explanation in artificial intelligence: Insights from the social sciences", *Artificial Intelligence*, 267 (Feb. 2019), pp. 1-38.

Moraschini, Tommaso (2023), "The Algebra of Logic", Lecture Notes.

Nakov, Georgi and Davide Emilio Quadrellaro (2022), *Algebraizable Weak Logics*.

Nolan, Daniel (1997), "Impossible Worlds: A Modest Approach", *Notre Dame Journal of Formal Logic*, 38, 4 (Oct. 1997).

Nute, Donald (1975), "Counterfactuals and the Similarity of Worlds", *Journal of Philosophy*, 72, 21, pp. 773-778.

— (1976), "David Lewis and the Analysis of Counterfactuals", *Noûs*, 10, 3 (Sept. 1976), p. 355.

Ono, Hiroakira (2019), *Proof Theory and Algebra in Logic*, Springer Singapore.

Over, David, Constantinos Hadjichristidis, Jonathan Evans, Simon Handley, and Steven Sloman (2007), "The probability of causal conditionals", *Cognitive Psychology*, 54 (Mar. 2007), pp. 62-97.

Palmer, Frank Robert (1986), *Mood and Modality*, Cambridge University Press.

Pearl, Judea (1988), *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann Publishers Inc.

— (2000), *Causality*, Cambridge University Press, Cambridge.

— (2017), "Physical and Metaphysical Counterfactuals: Evaluating Disjunctive Actions", *Journal of Causal Inference*, 5, 2, pp. 1-10.

Pollock, John L. (1976), *Subjunctive Reasoning*, Springer Netherlands.

Punčochář, Vıit (2021), "Inquisitive Heyting Algebras", *Studia Logica*, 109, 5 (Feb. 2021), pp. 995-1017.

Quadrellaro, Davide Emilio (2022), "On intermediate inquisitive and dependence logics: An algebraic study", *Annals of Pure and Applied Logic*, 173, 10 (Dec. 2022), p. 103143.

Raidl, Eric (2020), "Definable Conditionals", *Topoi*, 40, 1 (June 2020), pp. 87-105.

Ramsey, Frank Plumpton (1931), *The Foundations of Mathematics and Other Logical Essays. Edited by R.B. Braithwaite, with a Pref. By G.E. Moore. –*, London, England: Routledge and Kegan Paul.

Resconi, Germano, George J. Klir, and Ute St. Clair (1992), "Hierarchical uncertainty metatheory based upon modal logic", *International Journal of General Systems*, 21, 1 (June 1992), pp. 23-50.

Santorio, Paolo (2019), "Interventions in Premise Semantics", *Philosophers' Imprint*, 19.

— (forthcoming), "Probabilities of Counterfactuals are Counterfactual Probabilities", *Journal of Philosophy*.

Sartorio, Carolina (2006), "Disjunctive Causes", *Journal of Philosophy*, 103, 10, pp. 521-538.

Schulz, K. (2014), "Fake Tense in conditional sentences: a modal approach", *Natural Language Semantics*, 22, 2 (Mar. 2014), pp. 117-144.

Schulz, Katrin (2007), *Minimal Models in Semantics and Pragmatics: Free Choice, Exhaustivity, and Conditionals*, PhD thesis, University of Amsterdam: Institute for Logic, Language and Computation.

Schulz, Moritz (2017), *Counterfactuals and Probability*, Oxford University Press.

Segerberg, Krister (1989), "Notes on conditional logic", *Studia Logica*, 48, 2 (June 1989), pp. 157-168.

Sendłak, Maciej (2021), "Counterpossibles, story prefix and trivialism", *Synthese*, 199, 3-4 (Mar. 2021), pp. 7283-7301.

Shafer, Glenn (1976), *A Mathematical Theory of Evidence*, Princeton University Press.

Skyrms, Brian (1974), "Contraposition of the conditional", *Philosophical Studies*, 26, 2 (Oct. 1974), pp. 145-147.

— (1980a), *Causal Necessity*, Yale University Press, Princeton.

— (1980b), "The Prior Propensity Account of Subjunctive Conditionals", in *IFS*, Springer Netherlands, pp. 259-265.

Smets, Sonja and Fernando Velázquez-Quesada (2022), "Philosophical Aspects of Multi-Modal Logic", in *The Stanford Encyclopedia of Philosophy*, ed. by Edward N. Zalta and Uri Nodelman, Winter 2022 edition, Metaphysics Research Lab, Stanford University.

Stalnaker, Robert (1968), "A Theory of Conditionals", in *Studies in Logical Theory*, ed. by Nicholas Rescher, American Philosophical Quarterly Monograph Series, 2, Blackwell, Oxford, pp. 98-112.

— (1970), "Probability and Conditionals", *Philosophy of Science*, 37, pp. 64-80.

— (1975), "Indicative Conditionals", *Philosophia*, 5, 3, pp. 269-286.

— (1980), "A Defense of Conditional Excluded Middle", in *IFS*, Springer Netherlands, pp. 87-104.

Stalnaker, Robert C. and Richmond H. Thomason (2008), "A semantic analysis of conditional logic1", *Theoria*, 36, 1 (Feb. 2008), pp. 23-42.

Starr, W. (2022), "Counterfactuals", in *The Stanford Encyclopedia of Philosophy*, ed. by Edward N. Zalta and Uri Nodelman, Winter 2022 edition, Metaphysics Research Lab, Stanford University.

Starr, William B. (2013), "A Uniform Theory of Conditionals", *Journal of Philosophical Logic*, 43, 6 (Sept. 2013), pp. 1019-1064.

Van Benthem, Johan (2018), "Implicit and Explicit Stances in Logic", *Journal of Philosophical Logic*, 48, 3 (Nov. 2018), pp. 571-601.

Van Fraassen, Bas C. (1974), "Hidden Variables in Conditional Logic", *Theoria*, 40, 3, p. 176.

— (1976), "Probabilities of Conditionals", in *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*, Springer Netherlands, pp. 261-308.

Warmbrod, Ken (1982), "A defense of the Limit Assumption", *Philosophical Studies*, 42, 1 (July 1982), pp. 53-66.

Weiss, Yale (2019), *Frontiers of Conditional Logic*, PhD thesis, City University of New York (CUNY).

Wen, Xuefeng (2021), "Modal Logic via Global Consequence", in *Logic, Language, Information, and Computation*, Springer International Publishing, pp. 117-132.

Williamson, Timothy (2010), "Modal Logic within Counterfactual Logic", in *Modality*, Oxford University PressOxford, pp. 81-96.

Woodward, James (2003), *Making Things Happen: A Theory of Causal Explanation*, Oxford University Press, Oxford.

List of Figures

List of Tables