2024-01-26

Counterfactual Logic: A Modern Overview

Rios Flores, Mohamar

UNIVERSITY OF CALGARY

Counterfactual Logic: A Modern Overview

by

Mohamar Rios Flores

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN MATHEMATICS AND STATISTICS

CALGARY, ALBERTA

JANUARY, 2024

Abstract

Lewis (1973) described a family of logics for reasoning about counterfactual statements. These logics also contained additional connectives for reasoning about comparative possibility statements and modalities. Moreover, Lewis described a possible world semantics involving a "sphere system" that effectively lets you talk about some worlds being more "similar" to a given world than others. The resulting theory is very powerful and flexible in many ways similar to the theory of normal modal logics.

Unfortunately, Lewis provided an extremely terse formal description of the theory with many important theorems, definitions, and details omitted or described vaguely. There are no other resources that present a complete formal description of this theory. Without this formal description one is unable to propose new logics in this family and prove soundness and completeness theorems for them.

To remedy this, we present a reformulation of Lewis' theory using modern methods and notation. All definitions are described formally and all core results are proven in explicit detail. We include an informal overview of the different kinds of statements associated with each connective. The semantics is reformulated in terms of frames. The syntax includes additional rules and theorems to make it easier to use. Both the soundness and the strong soundness theorems are proven. The completeness theorem uses canonical models and the entire construction is shown explicitly. The first edition of *Counterfactuals* (Lewis, 1973) had an error exposed by Krabbe (1978) that led to the second edition (Lewis, 2001) having a more complex definition. So we present a reformulation of Krabbe's construction within our modernized setting.

Preface

This thesis is original, unpublished, independent work by the author, Mohamar Rios Flores.

Acknowledgments

I'm deeply thankful to Kristine for taking me on as her student and welcoming me into her group of grad students. I'll always look back fondly on the pastry meetups our group had in the math lounge. I'm grateful for all the academic and career advice I've received from Kristine over the years, it's evident to me that she is a person who really cares about the development of her students.

When I began my research project with Kristine I did not expect that my work would come to involve so many intricate and nuanced logical matters. I feel both very thankful and fortunate that Richard was willing to take the time to advise me on my project, which he did a lot of before he became my co-supervisor. I truly do not know if I would've been able to finish this project without his support.

To Dr. Clifton Cunningham and Dr. Gillman Payette, who took the time to read through my thesis and provide feedback as members of my committee, I offer my sincere appreciation.

I wish to thank my friend, Erik Chan, for taking the time to revise a draft of this thesis and providing a lot of useful feedback.

Last but not least, I give thanks to my friends and family, without whom I would not be who I am. To those in far away places whom I have not seen in years, I carry you in my heart wherever I go. To those nearby whom I don't see often enough, I'm glad I've come to know you and I'm grateful for the time we spend together. Thank you all for always supporting me.

To my younger self who as an undocumented immigrant never believed in a possible future where I would attend graduate school.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

In 1973, David Lewis published the book, *Counterfactuals* (Lewis, 1973), where he developed a theory of counterfactual logic using what he called "variably strict conditionals" and a semantics involving what he called "systems of spheres." Lewis also published a number of related papers in the same time period (Lewis, 1971, 1981a,b) that offered further insights into certain counterfactual logics or described alternate semantics. The book remains, to this day, not only the the primary resource on Lewis' semantics but one of the only resources that discusses his theory in full generality. This is partly because Lewis' theory actually yielded a large family of logics and many texts referencing his work are only interested in specific ones. The first edition of *Counterfactuals* contained an error in a key definition that was addressed in a paper by Krabbe (1978) and subsequently corrected for the second edition (Lewis, 2001).

There existed other approaches to counterfactuals prior to Lewis' publishing *Counterfactuals* that used different syntax and semantics. Lewis included a chapter in *Counterfactuals* (Lewis, 2001, Chapter 2) where he discusses various other contemporary semantics as reformulations and/or special cases of his approach. Stalnaker and Thomason's approaches (Stalnaker and Thomason, 1970; Stalnaker, 1968) were addressed both in *Counterfactuals* and in a subsequent paper by Lewis (1971). For an alternate perspective one can refer to (Sider, 2010, Chapter 8) which compares a reformulation of Stalnaker's theory with a reformulation of Lewis' semantics. The vast majority of approaches to counterfactual logic have relied on the theory of conditional logics for their syntax (Segerberg, 1989; Chellas, 1975). For a modern general overview of conditional logics one may refer to Weiss (2019). A conditional logic reformulation of two of the logics described by Lewis' can be found in Unterhuber (2013) with a discussion about how they relate to other conditional logics like (Segerberg, 1989; Chellas, 1975). For a big overview of how different semantics for conditional logics are related (including various formulations of semantics for counterfactuals) one may refer to Nute (1980). There have also been some successful non-logical approaches to reasoning about counterfactuals like Judea Pearl's (Pearl and Mackenzie, 2018) structural causal model that focuses on analyzing causality rather than tautologies.

Despite the popularity of the conditional logic approaches, the Lewis' approach continues to be well regarded as one of the classical approaches. That said, Lewis' *Counterfactuals* had a lot of innovative ideas when it was published and, understandably, it spends the vast majority of its pages attempting to build intuition for the readers. Unfortunately, this means that the actual formal description of the logic (i.e., all of the definitions and theorems you would expect from a logic textbook) is relegated to the very short final chapter of the book. Within that chapter he defines the syntax, semantics, gives a long list of axioms and properties (resulting in a family of 26 logics), proves soundness and completeness (with non-trivial canonical model constructions) in a span of merely twelve pages. He also discusses axiomatizations for specific logics, derived modal sub-logics, and other material for the remaining pages. To say that the material is terse and that important details are omitted is an understatement. Many statements are given without proof,[1] many technical definitions and theorem statements are stated vaguely in prose and at times merely alluded to.

There do not exist any other textbooks or materials (modern or otherwise) that explicitly present the material in that chapter. This thesis will not only

[1]This was the case for a false statement in the first edition (Lewis, 1973) that was key to proving completeness.

explicitly define and prove all of the core results in that chapter[2] but to reformulate the definitions and theorems in the style of modern formal logic textbooks with the addition of concepts like frames, valuations over propositions,[3] and canonical logics. I'm basing a lot of the reformulations on modern modal logic texts, primarily Zach's *Boxes and Diamonds* (Zach, 2019). I won't explicitly prove results about each of Lewis' 26 logics as that would be impractical, instead my goal is to develop all of the core theory necessary so that a reader can create their own counterfactual logic.

I initially set out to solve a different problem when I started my thesis project. Lewis' semantics use a possible world approach where you have a set of worlds and a mathematical structure on them called a system of spheres. This system of spheres is used to develop a notion of similarity between worlds (where some worlds are more similar to a given world than others). However, this structure is very generally defined and while that offers a great deal of flexibility it can also make it unwieldy. My initial goal was to attempt to find a subfamily of conditional logics whose sphere systems could be reformulated as topologies in a logically meaningful way (i.e., to develop topological models for a subfamily of counterfactual logics) and then to study the interplay between logic and topology. However, I found it very difficult to make technical progress on the problem due to many unclear definitions and theorem statements in *Counterfactuals* as well as general a lack of secondary resources. Though I did ultimately make some headway I quickly realized that I lacked the formal machinery needed in order to precisely prove the things I needed to prove. The only way forward was to reformulate and prove everything myself. What started as a small problem has now become my thesis.

Dear reader, if you are reading this because you found yourself in my unfortunate shoes then please know that, though separated by time and space, you are not alone and I hope you find this text useful. If you are reading this for any reason, then I really hope you enjoy it.

[2]Aside from the derived modal logic and decidability results which we leave for future work.

[3]Unterhuber (2013) provides definitions of frames and models for Lewis' "sphere system" semantics in a modern style but he does not prove any of the necessary theorems and just refers to *Counterfactuals*.

1.1 Prerequisites

We assume that the reader has a basic understanding of (naive) set theory and a basic understanding of classical propositional logic. Having some exposure to modal logic and relational models may be helpful but not required.

1.2 What are Counterfactuals?

Before we begin discussing the 'how' of counterfactual logic we should discuss the 'what' and the 'why.' As such, we'll begin by informally explaining and giving examples of some of the sorts of statements we'll be interested in.

First, recall that in classical logic we have conditional statements of the form:

> "If A, then B."

These statements have a property called **vacuous truth** where if A is false then we conclude that the sentence is true.

For example, consider the statement:

> "If it's raining, then the ground is wet."

Note that there are several reasons the ground could be wet, other than rain. So, if we walk outside and it's not raining but the ground is wet does that mean the statement is false? No, we simply default to saying that the statement is true (because it's not a lie) even though it doesn't tell us anything useful. In such cases we say that the sentence is vacuously true.

However, when you step back and think about the sort of if-then questions that people ask you'll find many real-world examples where vacuous truth is not at all the answer that people are interested in (or even aware of). It seems, that in some sense, there are perhaps different types of if-then questions that people are asking. A counterfactual statement is a prime example of this.

A counterfactual[4] is a statement of the form:

[4] We'll often refer to a "counterfactual statement" as simply a "counterfactual."

> "If it *were* the case that A, then it *would be* the case that B."

These statements are called counterfactuals because they are typically making a claim "against the facts," hence "counterfactual."

For example, consider the statement:

> "If it *were* the case that I knew how to play guitar, then it *would be* the case that I would be popular at parties."

Now, suppose that the person making this statement does not know how to play guitar. It's very hard to argue that this person is making an extraneous vacuously true statement. Consider, for example, if you were to respond to them with "that's a true statement," how would they interpret it? Most likely that they would think you imagined scenarios of what the world would look like if they knew how to play guitar and made the judgement that in those scenarios they are popular.

Another way to phrase this, in terms of "possible worlds," is to say that you imagined the most similar set of worlds[5] where they know how to play guitar and made the judgement that in those worlds they are popular. Lewis' theory of counterfactual logics uses possible world semantics and this is often how we'll be looking at things going forward. It's assumed that some worlds are more similar to ours than others and that some worlds could even be equally similar to ours.

This guitar example may not seem hugely important, but there are in fact a great many important questions that people encounter on a daily basis where some level of counterfactual reasoning is required, such as:

1. a healthcare worker trying to reason about the treatments they could be giving or could have given,

2. an electricity company trying to reason about the effect the weather is having on their green energy production and prices,

[5]We're performing a slight simplification here. We'll see later that, for technical reasons, there doesn't always exist a 'set of the most similar worlds' where a given statement holds but this is still the key underlying idea for our definitions later.

3. a person who just had a fight with their partner and is trying to reason about different choices they could've made.

An abundance of questions like this is the reason why there have been so many approaches towards reasoning about counterfactual statements over the years.

At first glance it may look like all counterfactuals look and behave the same way. However, there seem to actually be many different kinds of counterfactual questions that people may have.

For example, consider the following two statements:

> "If it *were* the case that I worked hard every day, then it *would be* the case that I would get a raise."

> "In a perfect world, if it *were* the case that I worked hard every day, then it *would be* the case that I would get a raise."

Suppose that the person asking these questions does work hard every day but has not gotten a raise. They would probably evaluate the first sentence as false, because they do work hard every day and despite that they haven't gotten a raise. In some sense, when making a judgement about the first sentence, the most similar worlds where they work hard every day includes their own world. However, things are different for the second statement. Here, a restriction is placed on the sorts of similar worlds one can consider when making a judgement so that we only consider perfect worlds (this most certainly excludes the world that person is in). I'd argue that they would probably evaluate the second statement as true — because it's hard to imagine a perfect world where someone doesn't get a raise for working hard.

There are other flavors of counterfactual statements one could consider. Statements about the future, the past, worlds with different laws of physics, and so on. We can also obtain different flavors of counterfactual statements by changing what it means for another world to be considered similar to ours. For example, maybe we want to consider another world to be similar to ours if:

1. it has a lot of the same facts as ours, or
2. it is possible and easy to change the facts of our world to match theirs, or
3. it is a possible near-future version of our world.

One thing to note is that when we evaluate a counterfactual to be true at our world, it doesn't just tell us things about other worlds, it also tells us things about the relationship between our world and other worlds. A useful way to make use of this is by changing what we mean when we say the term "worlds." Some examples are:

1. Each city could be considered to be a different "world."
    a) We can say that another "world" is similar to "our world" if it is cheap and easy for us to travel there.
    b) We can evaluate sentences like:

        > "If it *were* the case that gambling is legal, then it *would be* the case that there is a giant LED sphere nearby."

        If this statement were evaluated as true, then it would mean that the cheapest and easiest place for one to travel to for gambling is Las Vegas. Though to be clear, depending on one's circumstances this statement might not be true.[6]
2. Every person has developed a model of the world, based on their experiences and perceptions, that they carry around in their head, and in that sense we may consider each person as a different "world."[7]
    a) We can say that another person's "world" is similar to "our world" if we find it easy to relate to them and see things from their point of view.

[6]At the time of publishing Las Vegas is the only place on Earth with a giant LED sphere (which is simply named 'Sphere').

[7]Consider Jakob von Uexküll's notion of an umwelt.

b) We can evaluate sentences like:

> "If it *were* the case that the Earth is flat, then it *would be* the case that the government is brainwashing people with radio waves."

If this statement were evaluated as true, then it would mean that the most relatable people to you, who believe that the Earth is flat, also believe that the government is brainwashing people with radio waves.

3. We may consider each 256×256 pixel RGB picture as a "world."

a) We can say that another "world" is similar to "our world" if it's possible to convert "our world" into their "world" by subtly perturbing the pixels of "our world."

b) Given a machine learning image classifier, we can evaluate sentences like:

> "If it *were* the case that the image classifier detects a dog, then it *would be* the case that the image classifier detects a ball."

If this statement is evaluated as true then it might mean that the dataset for our image classifier was biased and didn't contain sufficient pictures of dogs without balls and therefore the classifier has concluded that "a dog without a ball is not a dog at all." It could also mean that we started with a picture with a ball.[8,9]

[8]There are real-world applications of counterfactuals in machine learning interpretability. One strategy involves searching for antecedents that would make the consequent in our counterfactual true. For our example, this would mean finding answers to the question "What is the simplest alteration I can make to my picture to guarantee that it includes a ball." For a more practical example, we could look at a model that assesses the credit risk of applicants. Then the question becomes "What are the most minor changes that a given applicant would have to make to their application in order to be deemed likely of having a good credit risk," (see Molnar, 2020, Example 9.3.2).

[9]A closely related notion to counterfactuals used in machine learning interpretability are anchors, which are kind of the opposite notion to counterfactuals. See Ribeiro et al. (2018) for a similar real-world application using anchors.

In each of these cases, it's clear that by evaluating a counterfactual from the perspective of a "world," we actually end up learning information about how that "world" is related to other "worlds."

There's one more way in which counterfactual statements may vary. Consider the sentence:

> "If it *were* the case that dogs originated from the Sun, then it *would be* the case that pigs fly."

Suppose that there are no similar worlds to ours (at all) such that dogs originated from the Sun. How should we evaluate this sentence? There are two ways we could do things.

1. We could evaluate it as vacuously true.

2. We could evaluate it as false.

We could consider these to be two different kinds of counterfactual statements. We regard the second one as a "strong counterfactual"[10] statement because it does not allow any notion of vacuous truth. As we'll see later, counterfactual logic actually provides us with a connective for each of these so we aren't forced to choose.

The reason for providing many different examples of different flavors of counterfactuals is to try to convey how flexible and powerful these notions can be with a little imagination. Lewis described a family of 26 counterfactual logics and counterfactuals behave differently in each of those logics. Don't think of counterfactual logic as a single tool for a single job, instead think of it the way that people think of modal logic, as a large family of tools for a large family of jobs.

As a quick summary, there are many different flavors of counterfactuals. Some ways that they vary is by having a different:

1. restriction on the sets of worlds.

2. meaning of "similarity" between one world and another.

[10] This terminology comes from Lewis (2001).

3. meaning of "worlds."

4. behavior when there are no similar worlds that satisfy the antecedent of the counterfactual (vacuously true vs false).

It may not be clear now, but the first three items on this list are very similar. Later, once we have developed our semantics formally, we will see that the first three items will be addressed in the same simple way. The fourth item on this list isn't a choice we have to make, we can reason about both types of counterfactuals together.

We do not have a maneuver analogous to the contrapositive for counterfactual statements. For example, suppose the following counterfactual statement is true:

> "If it *were* the case that the bus route goes by my house, then it *would be* the case that I'd ride the bus."

Performing a maneuver analogous to taking a contrapositive would yield this statement:

> "If it *were* the case that I would not ride the bus, then it *would be* the case that the bus route does not go by my house."

The first statement only tells us that in the most similar worlds if the bus route goes by their house then they would ride the bus. However, there could be less similar worlds where they do not ride the bus for some other reason despite the bus route still going by their house. In such cases the second statement would be false.

We'll briefly discuss the closely related dual counterfactual statement:

> "If it *were* the case that A, then it *might be* the case that $\boldsymbol{B}$."

This type of statement is related to but functions slightly differently to the counterfactual statements we've discussed so far.

For example, consider the statement:

> "If it *were* the case that I knew how to play guitar, then it *might be* the case that I would be popular at parties."

This statement says that in the most similar worlds where this person knew how to play guitar, there is at least one world where they are also popular at parties.

To see the relationship between counterfactual statements and dual counterfactual statements, consider the following two statements:

> "If it *were* the case that A, then it *would be* the case that B **is false**."

> "If it *were* the case that A, then it *might be* the case that B."

Note that the consequent in the first statement is false. The second statement is true if and only if the first statement is false.

Both the counterfactual and the strong counterfactual (which does not allow any notion of vacuous truth) have duals and they behave differently when it comes to vacuous truth. The dual counterfactual does not allow any notion of vacuous truth, but the dual strong counterfactual does. In some sense, the dual of the strong counterfactual is actually weaker. In counterfactual logic we'll have separate connectives for each of these.

1.3 What is Comparative Possibility?

Another type of statement that we may be interested in studying is one in which we compare the possibility of two things. Such a statement may look like this:

> "It is *at least as possible* that A as it is that B."

We refer to statements that compare possibility like this as comparative possibility statements. We must also stress that though we often compare the *probabilities* of things, this is not at all the same thing as comparing *possibility*. We'll clarify what we mean by this as we continue.

For example, consider the statement:

> "It is *at least as possible* that it is raining as it is that I've been struck by lightning."

Suppose the person making this claim is in an area where dry thunderstorms[11] are rare but not impossible. This statement does not say that if they've been struck by lightning then it is raining. The way it's written it doesn't rule out the possibility of being struck by lightning without rain. Consider, for example, if you were to respond to them with "that's a true statement," how would they interpret it? They would probably think you imagined several possible scenarios of what the world would look like in either case and concluded that there are scenarios where it is raining that are at least as similar to our world as the most similar scenarios where they've been struck by lightning.

Another way to phrase this in terms of "possible worlds" is to say that you imagined the most similar set of worlds where they've been struck by lightning, and found that that this set included worlds where it is raining. Note that this is fundamentally different to how we compare probabilities. For one, probability is quantifiable but possibility is not (i.e., we cannot compute the possibility that it is raining). For another, if something is only true in a few very similar worlds to ours but is false in all other worlds, then we would still consider it as more possible than many other things that are true in far less similar worlds.

We can also try to reason about the negation of this statement. To do so, consider what it would have to mean for the statement to be false. That is, one would have to claim that the possibility that they've been struck by lightning is greater than the possibility that it is raining. Rephrasing this we have the comparative possibility statement:

> "It is *more possible* that I've been struck by lightning than that it is raining."

Again, supposing that the person saying this is in an area where dry thunderstorms are rare but not impossible leads us to conclude that this statement is false. For it to be true we would have to show that the most similar set of worlds containing worlds where they've been struck by lightning do not include any worlds where it is raining.

[11]These are a rare kind of thunderstorm where rain evaporates before it hits the ground.

When we presented our first example statement we mentioned that it does not say that "if they've been struck by lightning then it is raining." Let's briefly compare our first example of a comparative possibility statement to the conditional:

> "If I've been struck by lightning, then it is raining."

This is in one respect a vacuously true statement about our world, and in an another respect, if we assume this statement is true at all worlds, it is a significantly stronger statement. Taking the contrapositive yields an equivalent statement:

> "If it is not raining, then I have not been struck by lightning."

Neither of these statements capture the meaning of our comparative possibility statement. However, the fact that we can take contrapositives of conditionals raises the question of whether or not it makes sense to perform a similar maneuver on comparative possibility statements.

For example, consider the statement:

> "It is *at least as possible* that I haven't been fired as it is that Earth hasn't stopped turning."

Suppose the person making this statement currently has a job and the Earth is turning. Both of these statements are true in the most similar world to the one belonging to this person. Then, the statement is true, because in the most similar world where the Earth hasn't stopped turning it is also the case that they haven't been fired.[12]

Let's now attempt to perform a maneuver analogous to taking a contrapositive for this statement.

> "It is *at least as possible* that Earth has stopped turning as it is that I have been fired."

[12]This is a statement where our intuition about probability may attempt to lead us astray. One may be tempted to (incorrectly) say that the statement is false because the probability that the Earth will stop spinning is much lower than the probability that they won't get fired.

Arguably it is impossible for the earth to stop turning but it is possible that they have been fired. Therefore, it is clearly more possible that this person has been fired than that the world has stopped turning. To put it another way, the set of the most similar worlds containing worlds where this person has been fired does not contain any where the world has stopped turning. Therefore another difference between the conditional and comparative possibility statements is that we can take contrapositives with conditionals but we can't perform an analogous maneuver with comparative possibility statements.

Speaking of impossible sentences, it's clear that any sentence is at least as possible as any impossible sentence, such as "The Earth has stopped turning."

One property that we do have is a form of transitivity. Suppose the following two statements are true:

"It is *at least as possible* that A as it is that B."
"It is *at least as possible* that B as it is that C."

Then it must be the case that the following statement is also true:

"It is *at least as possible* that A as it is that C."

We can't formally prove it at this point but the idea is that the most similar set of worlds containing a world where C is true also contain a world where B is true, and the most similar set of worlds containing a world where B is true also contain a world where A is true. Therefore, the most similar set of worlds containing a world where C is true also contains a world where A is true.

We are also able to make comparative possibility statements about things being equally possible:

"It is *equally possible* that A and that B."

This is equivalent to both of the following statements being true simultaneously:

"It is *at least as possible* that A as it is that B."
"It is *at least as possible* that B as it is that A."

Note that in order for two things to be equally possible it does not mean that they are true in the same world. You could be standing in line to receive one of several possible prizes and from your perspective each one could be equally possible (even though you only get one prize).

Like with counterfactuals, we can add restrictions to worlds, change what we mean by similarity, or change what we mean by worlds. Some of the examples we discussed for counterfactuals can be adapted for comparative possibility statements.

You may have noticed that while comparative possibility and counterfactuals say different things, they seem to deal with very similar reasoning. This is not a coincidence. In fact, counterfactuals can be defined in terms of comparative possibility statements and vice versa. We'll use this fact later to define everything in terms of comparative possibility when we begin to formally describe things in the next chapter. We'll use comparative possibility because it will make some definitions, proofs, and statements simpler.

1.4 What are Modalities?

The last type of statement that we'll be informally explaining are modalities. For readers with some background in modal logic, the modalities here work similar to how they work in modal logic. Other readers do not need to be worried, we will go over all you need to know.

Modalities come in pairs that are dual to each other. In counterfactual logic we will encounter two pairs of modalities, one pair will be called "outer modalities" and another one will be "inner modalities." The outer modalities deal with accessible worlds. So, before we move on we need to address this notion.

One thing we've glossed over is that while some worlds are more similar to ours than others, it's not necessarily the case that every world is similar to ours by some extent. Some worlds may not be in any way similar to ours at all and we will say that those are inaccessible. For instance, consider the case where similarity is defined such that a world is similar to ours if it is a possible near-future version of our world. Then, any world that's a version of the past or that is not a version of our world is not similar to our world at all and therefore "inaccessible." Conversely, the set of accessible worlds will refer to all of the worlds similar to ours.

We can now begin discussing our outer modalities. We'll start by showing how a statement using an outer modality may look:

> "It is *necessarily true* that A."

Though we write "necessarily true" here, we will primarily refer to these modalities with notation later on. This is because, depending on how similarity works between worlds, other terms may be more appropriate.

Consider the example:

> "It is *necessarily true* that smoking causes cancer."

To say that a statement is necessarily true, at our world, means that it is true at every accessible world. We do not consider inaccessible worlds because they are irrelevant from the perspective of our world. So, for this statement to be true it would mean that at every world similar to ours, in any way, it must be the case that smoking causes cancer. Though, there could still be inaccessible worlds where smoking does not cause cancer.

For our dual outer modality, we have a statement of the form:

> "It is *possibly true* that A."

As was the case with our other outer modality, we use the term "possibly true" here, but in some cases other terms may be more appropriate.

Consider the example:

> "It is *possibly true* that there is a cure for cancer."

Similar to our last modality, to say that a sentence is possibly true, at our world, means that it is true in at least one accessible world. In order for this statement to be true it would mean that there exists a world, similar to ours, where there is a cure for cancer.

Next, we will briefly discuss the inner modalities. Instead of accessible worlds, these refer to the set of most similar worlds. We could call these the "innermost accessible worlds." If we've decided that the only innermost accessible world from ours is our own then the inner modalities don't say anything interesting, but if we've decided that the innermost accessible worlds contain many other worlds (possibly including our own) then they have more interesting behavior very similar to the outer modalities.

Let's start describing how inner modalities work by showing how one might look:

> "It is *immediately necessarily true* that A."

Though we write "immediately necessarily true" here, we'll always write it with notation after the next section. Lewis' doesn't ever refer to sentences with inner modalities in natural language, so the term "immediately necessarily true" is one I'm using for lack off a better term.

Consider the example:

> "It is *immediately necessarily true* that the stock market hasn't crashed."

For this statement to be true it must be the case that the stock market hasn't crashed at every innermost accessible world. However, there could be worlds that are less similar and still accessible where the stock market has crashed.

For our dual inner modality, we have a statement of the form:

> "It is *immediately possibly true* that A."

As was the case with our other inner modality, the term "immediately possibly true," won't be used again after the next section as I'm only using it for lack of a better term.

Consider the example:

> "It is *immediately possibly true* that I have the flu."

For this statement to be true it must be the case that there is an innermost accessible world where I have the flu. It doesn't necessarily have to be our world, but it must be among the innermost accessible worlds.

Innermost and outermost modality pairs can both be defined in terms of either counterfactuals or comparative possibility statements, though the latter is easier to work with. On the other hand, we cannot define counterfactuals or comparative possibility statements in terms of any of our modalities.

1.5 What is Counterfactual Logic?

We won't formally discuss counterfactual logic until later chapters. For now, however, it may be useful to provide an informal overview of the main ideas. First of all, counterfactual logic is not just one logic. Rather, like "modal logic," the term actually refers to a large family of logics. Lewis explicitly listed 26, but more can be described. Second, counterfactual logic doesn't just let you reason about counterfactual statements — it also lets you reason about comparative possibility and two pairs of modalities, all at the same time. One could view counterfactual logic as a more fine grained version of modal logic, but this would be a large oversimplification.

As with many logics, we will discuss it in terms of both semantics and syntax. In the previous sections we spoke about how to evaluate statements by referring to worlds, similarity, and truth values. This is what semantics of the logic refers to, the way that we assign "meaning" to sentences. By adding properties to our semantics we can change their behavior and in doing so obtain different counterfactual logics. We'll focus more on the semantics side of things in these earlier chapters because they help us develop an intuition about how this family of logics behaves. On the syntax side of things we will describe how to construct formal logics, as axiomatic systems, using derivations. Later, in the chapter on syntax, we'll first develop one axiomatic system and then, by adding additional axioms, we will obtain other axiomatic systems.

In either case we'll have to go over the connectives. Counterfactual logic contains a large number of connectives, far more than modal logic:

1. The classical connectives:

   0-ary $\bot, \top$,

   1-ary $\neg$,

2-ary $\land, \lor, \rightarrow, \leftarrow, \leftrightarrow$,

2. The non-classical connectives:

1-ary $\Box, \Diamond, \boxdot, \diamonddot$,

2-ary $\preccurlyeq, \prec, \approx, \Box\!\!\Rightarrow, \Diamond\!\!\Rightarrow, \Box\!\!\rightarrow, \Diamond\!\!\rightarrow$.

As we've mentioned in previous sections, many of these are interdefinable, so even though it is a lot of connectives, rest assured that it is not as overwhelming as it looks. The classical connectives behave the same way they do in classical propositional logic. For the non-classical connectives we provide tables that list the connectives with the corresponding sentences we discussed in previous section. There are several types of counterfactuals which we describe in table 1.1 with the third column indicating whether it's a dual, strong, or dual strong version. The remaining non-classical connectives are described in table 1.2.

"If it *were* the case that A, then it *would be* the case that B."	$A \Box\!\!\rightarrow B$	
	$A \Box\!\!\Rightarrow B$	strong
"If it *were* the case that A, then it *might be* the case that B."	$A \Diamond\!\!\rightarrow B$	dual
	$A \Diamond\!\!\Rightarrow B$	dual strong

Table 1.1: The counterfactual and dual counterfactual connectives.

"It is *at least as possible* that A as it is that B."	$A \preccurlyeq B$
"It is *more possible* that A than that B."	$A \prec B$
"It is *equally possible* that A and that B."	$A \approx B$
"It is *necessarily true* that A."	$\Box A$
"It is *possibly true* that A."	$\Diamond A$
"it is *immediately necessarily true* that A."	$\boxdot A$
"it is *immediately possibly true* that A."	$\diamonddot A$

Table 1.2: The comparative possibility and modality connectives.

For now we'll avoid further discussion on syntax as those notions are all very formal. We will, however, give an informal example of a sphere model in order to motivate the formal description of our semantics in the following chapter.

1.6 An Informal Introduction to Sphere Models

Our semantics will formalize the notions of worlds, similarity, and truth that we referred to in previous sections. It's easy to think about a set of worlds and some sentences being true at some worlds, however it is much more difficult to think about similarity. Afterall, what is it? Let's think about a very simple example. We'll suppose that we have a set of six worlds,

$$W = \{ w_0, w_1, w_2, w_3, w_4, w_5 \},$$

and that our world is w_0. We want our notion of similarity to be able to say that some worlds are more similar to ours than others. So, for instance, we want to be able to say things like that from the perspective of w_0, it is the case that w_0, w_1, and w_2 are more similar than w_3 and w_4. However, we also want to be able to say that w_0, w_1, and w_2 are equally similar to w_0, w_3 and w_4 are equally similar to w_0, and that w_5 isn't similar to w_0 at all. If we draw this it will look like fig. 1.1. Note that we don't draw w_5 because it's not similar at all. This diagram resembles two nested circles or *"spheres."* Arguably, the most

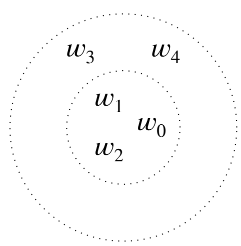


Figure 1.1: Drawing of what similarity may look like.

straightforward way to capture this idea is by first describing each of these "spheres" as a set as follows:

$$S_0 = \{ w_0, w_1, w_2 \},$$
$$S_1 = \{ w_0, w_1, w_2, w_3, w_4 \}.$$

Also, for purely technical reasons that will make sense later, we include $\varnothing$ as a sphere as well. We'll use the notation $\mathcal{O}(w_0)$ to refer to "the spheres around the world w_0" and we'll write it as follows:

$$\mathcal{O}(w_0) = \{ \varnothing, S_0, S_1 \} .$$

In other words,

$$\mathcal{O}(w_0) = \{ \varnothing, \{ w_0, w_1, w_2 \} , \{ w_0, w_1, w_2, w_3, w_4 \} \} .$$

Technically speaking, each world should have its own set of spheres around it (i.e., we should define $\mathcal{O}(w_1), \ldots, \mathcal{O}(w_5)$) in order for this to be a complete example. However, lets first look at some of our sentences from earlier and see how they would be interpreted in terms of spheres.

Recall our earlier counterfactual statement:

> "If it *were* the case that I knew how to play guitar, then it *would be* the case that I would be popular at parties."

We're going to imagine that we have six worlds, $\{ w_0, \ldots, w_5 \}$. Suppose that w_0 is our world and that these are the spheres around our world,

$$\mathcal{O}(w_0) = \{ \varnothing, \{ w_0 \} , \{ w_0, w_1 \} , \{ w_0, w_1, w_2, w_3, w_4, w_5 \} \} .$$

We won't worry about the spheres around other worlds yet. Suppose that we have the following sentences that we're interested in, G = "I knew how to play guitar," P = "I would be popular at parties," and F = "could fly." Next, lets suppose that:

1. G is true at worlds: $\{ w_2, w_4 \}$,
2. P is true at worlds: $\{ w_1, w_2, w_3 \}$, and
3. F is true at worlds: $\varnothing$.

Note that a sentence is either true or false at a world, therefore both G and P are false at our world, w_0.

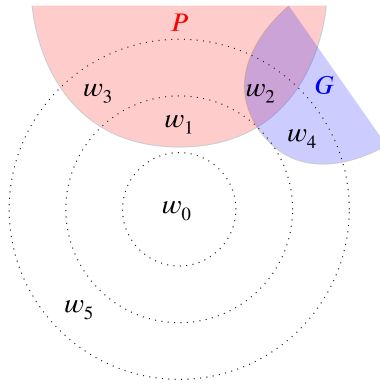


Figure 1.2: Drawing of $\mathcal{O}(w_0)$ with G and P highlighted. F is not depicted because there are no worlds where it is true.

We can draw all of this information as in fig. 1.2. Note that we shade in worlds where P and G are true and since F isn't true at any worlds then it does not appear in our diagram. Now we can ask evaluate which of the following statements are true or false from the perspective of our world, w_0:

1. $G \rightarrow P$: This statement is vacuously true, because G is false at w_0.

2. $G \;\Box\!\!\rightarrow P$: This statement is false because in the nearest sphere where G is true, it is not the case that $G \rightarrow P$ is true at every world within said sphere. In other words, the most similar worlds where G is true are w_2 and w_4 but while P is true at w_2, it is false at w_4.

3. $G \;\Box\!\!\Rightarrow P$: This statement is false for the exact same reason.

4. $G \;\Diamond\!\!\rightarrow P$: This statement is true because in the smallest sphere where G is true, there exists at least one world where both G and P are true, namely w_2.

5. $G \;\Diamond\!\!\Rightarrow P$: This statement is true for the exact same reason.

6. $F \;\Box\!\!\rightarrow P$: This statement is false, because there are no similar worlds where F is true.

7. $F \;\Box\!\!\Rightarrow P$: This statement is vacuously true, because there are no similar worlds where F is true.

8. $F \diamondto P$: This statement is false, because there are no similar worlds where F is true.

9. $F \Diamond\!\!\Rightarrow P$: This statement is vacuously true, because there are no similar worlds where F is true.

10. $G \preccurlyeq P$: This statement is false, because there exists a sphere containing a world where P is true that does not contain any worlds where G is true. Namely, the sphere $\{ w_0, w_1 \}$.

11. $P \prec G$: This statement is true, because there exists a sphere containing a world where P is true that does not contain any worlds where G is true. Namely, the sphere $\{ w_0, w_1 \}$. Compare this to the previous sentence, it is not a coincidence that the reasoning is the same.

12. $\neg P \prec G$: This statement is true, because there exists a sphere containing a world where P is false that does not contain any worlds where G is true. Namely, the sphere $\{ w_0 \}$.

13. $\Box G$: This statement is false, because there exist accessible worlds where G is false, namely w_0, w_1, w_3 and w_5.

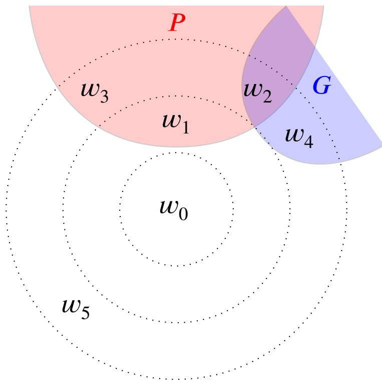
14. $\Diamond G$: This statement is true, because there exists at least one accessible worlds where G is true, namely w_0, w_1, w_3 and w_5.

15. $\Box \neg F$: This statement is true, because F is false at every world accessible from w_0.

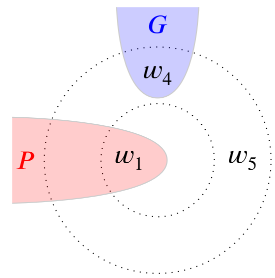
16. $\Diamond F$: This statement is false, because there do not exist any accessible worlds where F is true.

17. $\boxdot \neg P$: This statement is true, because P is false at every world in the smallest non-empty sphere, namely $\{ w_0 \}$.

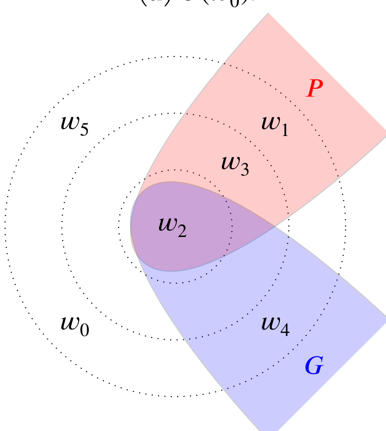
18. $\Diamond\!\!\!\cdot\; \neg P$: This statement is true, because there exists a world in the smallest non-empty sphere where P is false, namely at w_0.



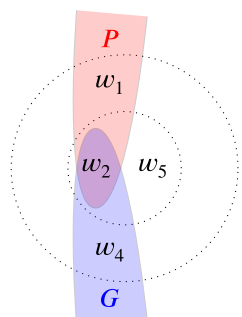
(a) $\mathcal{O}(w_0)$.



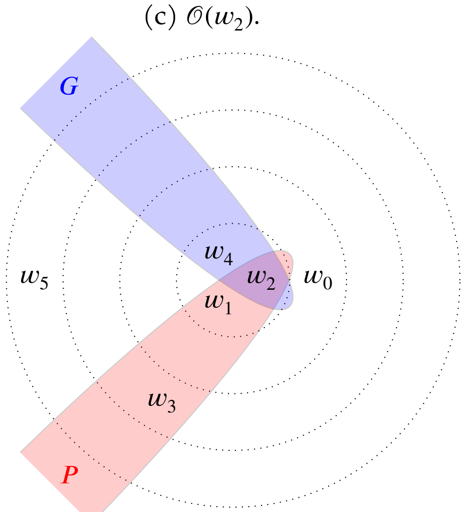
(b) $\mathcal{O}(w_1)$.



(c) $\mathcal{O}(w_2)$.



(d) $\mathcal{O}(w_3)$.



(e) $\mathcal{O}(w_4)$.

(f) $\mathcal{O}(w_5)$.

Figure 1.3: Drawing of the entire sphere system, $\mathcal{O}$.

It is important to keep in mind, however, that this example only tells us about whether or not these statements are true at w_0. If we want to know if they are true at w_1 then we must be able to refer to spheres around w_1, and so on.

Being able to say whether or not a statement is true at all worlds will actually be a very important thing later. Let's show what our example might look like if we defined the sets of spheres around the other worlds as well. This example will be counterintuitive on purpose in order to show the sorts of things that we allow. Let $\mathcal{O}$ be defined as follows:

$$\begin{aligned}
\mathcal{O}(w_0) &= \{\, \varnothing, \{\, w_0 \,\}, \{\, w_0, w_1 \,\}, \{\, w_0, w_1, w_2, w_3, w_4, w_5 \,\} \,\} \\
\mathcal{O}(w_1) &= \{\, \varnothing, \{\, w_1 \,\}, \{\, w_1, w_4, w_5 \,\} \,\} \\
\mathcal{O}(w_2) &= \{\, \varnothing, \{\, w_2 \,\}, \{\, w_2, w_3 \,\}, \{\, w_0, w_1, w_2, w_3, w_4, w_5 \,\} \,\} \\
\mathcal{O}(w_3) &= \{\, \varnothing, \{\, w_2, w_5 \,\}, \{\, w_1, w_2, w_4, w_5 \,\} \,\} \\
\mathcal{O}(w_4) &= \left\{ \begin{array}{l} \varnothing, \{\, w_1, w_2, w_4 \,\}, \{\, w_0, w_1, w_2, w_4 \,\}, \\ \{\, w_0, w_1, w_2, w_3, w_4 \,\}, \{\, w_0, w_1, w_2, w_3, w_4, w_5 \,\} \end{array} \right\} \\
\mathcal{O}(w_5) &= \{\, \varnothing \,\}.
\end{aligned}$$

It's helpful to have a drawing of all of these in order to get a sense of what we're talking about. For that we refer to fig. 1.3. Note that it is not necessary for $\boldsymbol{P}$ and $\boldsymbol{G}$ to be connected as in our drawing, we only shaded in this way to avoid having the image looking too cluttered. Some remarks about this example are:

1. While w_1 is similar to w_0 it is not the case that w_0 is similar to w_1. In general we may or may not have this sort of symmetry. Depending on how we're thinking about similarity it may not be desirable to to have symmetry, for instance if we're thinking in terms of possible future worlds.

2. The most similar world to w_0 is itself. This is also the case with w_1 and w_2. This property may sometimes be desired.

3. From the perspective of w_3, it is not the case that w_3 is accessible. This means that w_3 is not among the set of worlds that we consider similar

to w_3. Situations like this may be desired if we're interested in morally perfect worlds, for instance.

4. From the perspective of w_1, the set of accessible worlds does not include w_0 or w_3. This might be desirable in situations where some worlds are considered simply too different.

5. There are no worlds similar to w_5, this is pathological but there's no reason why it would be disallowed. Interestingly, our modalities exhibit some interesting edge case behavior in this case: $\Box G$ and $\Diamonddot G$ both become vacuously true at world w_5, while $\Diamond G$ and $\boxdot G$ both become false at world w_5. The reason for this comes from their formal description given in the next section (see theorem 2.3.5).

Many of the things we remarked on could be enforced at every world by adding properties to our model. For instance, if we wanted all of our worlds to behave like those mentioned in the second remark in our list, then we could add a property, called *centering*, that requires for any world, w, it must be the case that $\{ w \} \in \mathcal{O}(w)$.

This construction we've provided, consisting of a set of worlds, W, the function $\mathcal{O}$ describing spheres around worlds, and our assignment of truth values to sentences, are an example of a sphere model definition 2.3.2. Within our semantics they will be the primary objects of study. It is important to keep in mind that this sphere model was a simplified toy example to help build intuition. It's possible to have sphere models with an infinite number of worlds and in those cases we don't always have, for instance, a smallest sphere containing a world where a sentence is true (because we may have an infinite number of such spheres but no smallest sphere). We'll see later, when we formally describe how our connectives (see definition 2.3.3 and theorem 2.3.5) actually work with regards to our semantics, that the definitions look a little more complex but in spirit they will behave the same way that we've described (and if the set of worlds are finite they will behave exactly how we've described).

As a quick recap, in order to define a sphere model we only need these three things:

1. a set of worlds,
2. a set of spheres around each world,
3. for each sentence, a set of worlds where we'll assume it is true.

There are also some properties on the sets of spheres that are listed in the formal definition (see definition 2.2.1 and definition 2.3.2) but that we've tried to convey implicitly here.

1.7 Thesis Outline

The intent of this thesis is to provide all of the essential definitions and prove all of the essential theorems one would expect in a modern logic text on this topic. Broadly speaking, there are four primary goals we wish to cover:

1. Semantics: Develop sphere frames and models, truth and validity, and discuss various notions related to entailment.
2. Syntax: Develop axiomatic systems, various notions related to derivations, and maximally Σ-consistent sets of sentences.
3. Soundness: We prove two versions of the soundness theorem.
4. Completeness:
   a) We need to first describe canonical models, which are split off in their own chapter because they are highly non-trivial.
   b) Then we describe the conditions necessary for the completeness theorem to be true and we prove it.

The soundness theorem tells us that if something isn't a tautology in our semantics, then it is not a theorem in our syntax. In other words, the semantics are *sound* with respect to the syntax. The completeness theorem tells us that if something is a tautology in our semantics, then it is a theorem in our syntax. In other words, the semantics are *complete* with respect to the syntax. Together, soundness and completeness tell us that the semantics and syntax correspond with each other.

Semantics, Syntax, and Soundness each get one chapter. Completeness effectively gets two chapters, one for canonical models and one for completeness proper. Then there is an extra chapter discussing Krabbe's paper (which isn't required reading but is included for completeness). After that there's an additional chapter to hold additional theorems of results so as to not clutter the main text unnecessarily.

Regarding the chapter titled 'Krabbe's Counterexample', Krabbe (1978) exposed an error in the first edition of *Counterfactuals* (Lewis, 1973) that forced the second edition (Lewis, 2001) to change the definition of canonical models. Krabbe exposed this error by constructing a specific counterexample and we present a fully detailed modern reformulation of Krabbe's result.[13]

[13] This construction is also non-trivial and Krabbe's paper does it in under 3 pages by being incredibly terse and omitting a lot of technical details. As such I felt it would go against the spirit of this thesis to leave it out.

Chapter 2

Language and Semantics

We'll now begin to formally describe everything. We'll start by explicitly defining the underlying language for counterfactual logic. Next, we'll proceed to formally describe the semantics that were discussed in the introduction. Many definitions will be modified from Lewis (2001), primarily to modernize the theory and make it easier to work with from a formal standpoint. Our language will be simplified to a much smaller number of connectives which will make it so that proofs by induction over the language are much simpler. None of the theorems in this section are in Lewis' *Counterfactuals*, though a portion are actually definitions for Lewis.

2.1 Language

We'll start by developing the language we'll be working with throughout the rest of this thesis.

The Language L_V

A language is a set of strings over an alphabet (a set of symbols). To define our language we will first define an alphabet then give a grammar (i.e., set of rules) that tells us whether or not a string is a member of our language. Strings that are members of our language will be called sentences.

First we'll describe our alphabet.

Definition 2.1.1 (Alphabet for L_V)**.**
The alphabet for our logic is the set,

$$\mathscr{A}_V := \left\{ \mathbb{P}_n \;\middle|\; n \in \mathbb{N} \right\} \bigcup \{ (,), \bot, \to, \preccurlyeq \},$$

where

$$\textsc{Props} := \left\{ \mathbb{P}_n \;\middle|\; n \in \mathbb{N} \right\}$$

contains symbols called **propositions** or **atomic sentences** (both terms are commonly used interchangeably) and

$$\{ (,), \bot, \to, \preccurlyeq \}$$

contains symbols called **connectives**.

We've named the set Props because we will refer to it often.

Now we describe our language.

Definition 2.1.2 (Language, L_V, for **V**-logics)**.**
A **string** over the alphabet $\mathscr{A}_V$ is just a sequence of symbols in $\mathscr{A}_V$. We can now define our language, L_V as a set of strings over $\mathscr{A}_V$ satisfying the following recursive definition called a **grammar**.

L_V is the set of strings over $\mathscr{A}_V$ such that:

1. **Propositions:** Props $\subseteq L_V$
2. **Connectives:**
   a) **Bottom:** $\bot \in L_V$
   b) **Conditional:** If $A, B \in L_V$ then $(A \to B) \in L_V$
   c) **Precedes:** If $A, B \in L_V$ then $(A \preccurlyeq B) \in L_V$.

No additional strings are included in our language.

A string in L_V is called a **sentence in** L_V or simply a **sentence**.

Some remarks about our notation are in order.

Remark 2.1.3. We will be using capital letters like A, B, C to denote arbitrary sentences in our language L_V and we will use lowercase letters p, q, r to denote arbitrary propositions (i.e., elements of PROPS) in our language L_V.

Technically speaking, our definition of L_V is very rigid and excludes some strings that we might expect to be in L_V. It is helpful to view some examples. We will introduce some writing conventions later (see remark 2.1.7) to allow us to write sentences in L_V in a more relaxed way.

Example 2.1.4.
The following three strings are sentences inside L_V:

$$\mathbb{P}_2, \qquad (\mathbb{P}_3 \to \mathbb{P}_1), \qquad (\bot \preccurlyeq (\mathbb{P}_1 \to \bot)).$$

The following two strings are **not sentences** inside L_V:

$$\mathbb{P} \qquad (\bot \preccurlyeq \mathbb{P}_1 \to \bot).$$

The first one fails because every proposition must be enumerated by a natural number. The second one fails because a set of parenthesis is omitted (and worse, it is ambiguous as to where those parenthesis should go).

A useful detail worth pointing out is that our language is countable. We'll use this in a later proof to obtain an enumeration on our set of sentences.

Theorem 2.1.5.
The language L_V is countable.

Proof. The set of finite strings over a countable alphabet is countable. □

As you may have noticed, we are missing several familiar connectives which we would like to consider in our logic. We will implement these and others as *defined connectives* that will act as abbreviations for sentences in our language.

Definition 2.1.6 (Additional Connectives)**.**
Given sentences A and B in L_V, the following connectives:

$$\top, \neg, \wedge, \vee, \leftarrow, \leftrightarrow, \prec, \approx, \Box, \Diamond, \boxdot, \Diamond\!\!\!\cdot\,, \Box\!\!\Rightarrow, \Diamond\!\!\Rightarrow, \Box\!\!\rightarrow, \Diamond\!\!\rightarrow,$$

will be implemented by the *defined connectives* listed in table 2.1.[1]

Negation:	$\neg A := (A \rightarrow \bot)$
Top:	$\top := \neg\bot$
Conjunction:	$(A \wedge B) := \neg(A \rightarrow \neg B)$
Disjunction:	$(A \vee B) := (\neg A \rightarrow B)$
Reverse Conditional:	$(A \leftarrow B) := (B \rightarrow A)$
Biconditional:	$(A \leftrightarrow B) := ((A \rightarrow B) \wedge (A \leftarrow B))$
Properly Precedes:	$(A \prec B) := \neg(B \preccurlyeq A)$
Equipossible:	$(A \approx B) := ((A \preccurlyeq B) \wedge (B \preccurlyeq A))$
Outer Necessity:	$\Box A := (\bot \preccurlyeq \neg A)$
Outer Possibility:	$\Diamond A := (A \prec \bot)$
Inner Necessity:	$\boxdot A := (\top \prec \neg A)$
Inner Possibility:	$\diamonddot A := (A \preccurlyeq \top)$
Strong Counterfactual:	$(A \Box\!\!\Rightarrow B) := ((A \wedge B) \prec (A \wedge \neg B))$
Dual Strong Counterfactual:	$(A \Diamond\!\!\Rightarrow B) := ((A \wedge B) \preccurlyeq (A \wedge \neg B))$
Counterfactual:	$(A \Box\!\!\rightarrow B) := (\Diamond A \rightarrow (A \Box\!\!\Rightarrow B))$
Dual Counterfactual:	$(A \Diamond\!\!\rightarrow B) := (\Diamond A \wedge (A \Diamond\!\!\Rightarrow B))$

Table 2.1: Defined connectives

Defining our language with parenthesis in this way is a good way to eliminate ambiguity and confusion. However, it does lead to an excessive number of parenthesis which can also make sentences hard to read. One way to deal with this is by introducing writing conventions to reduce the number of parenthesis.

Some languages will define an order of operations (via a list of precedence rules) that describe a standardized method for parsing a sentence with missing parenthesis as one with parenthesis. A well-known example of this is the order of operations used for arithmetic (often described using acronyms

[1] Lewis defines $\Box A := (\bot \approx \neg A)$ and $\diamonddot A := (A \approx \top)$. We will see through theorems, theorem 2.3.7 and theorem 3.3.1, that our definitions are equivalent to Lewis' definitions but yield slightly simpler proofs.

like PEMDAS or BODMAS) which allow one to parse expressions, such as $3 - 4 \times 4 + 2$, in an equivalent unambiguous form, $(3 - ((4 \times 4) + 2))$. Classical propositional logic also traditionally uses an order of operations where some connectives are prioritized over others, however such an approach is impractical for us since we have seventeen connectives.

Instead we will adopt a fewer number of simpler conventions that will allow us to remove a few of the most unnecessary parenthesis.

Remark 2.1.7. For sentences in L_V:

1. We may omit the outermost pair of parenthesis in a sentence. For example, we may write $A \to (B \wedge C)$ and have it be understood as $(A \to (B \wedge C))$.

2. We may omit parenthesis involving multiple nested conjunctions. For instance, we may write $A \wedge B \wedge C \wedge D$ instead of $(((A \wedge B) \wedge C) \wedge D)$ or other variations like $((A \wedge B) \wedge (C \wedge D))$.

   This is allowed because conjunctions will behave here in the same way that they do in classical propositional logic.

3. We may omit parenthesis involving multiple nested disjunctions. For instance, we may write $A \vee B \vee C \vee D$ instead of $(((A \vee B) \vee C) \vee D)$ or other variations like $((A \vee B) \vee (C \vee D))$.

   This is allowed because disjunctions will behave here in the same way that they do in classical propositional logic.

Sentence Instances

Often we will want to use or prove a result about a collection of sentences that have the same structure. We can do this through sentence instantiation.[2] The basic idea is that by taking a sentence and only substituting propositions we'll end up with a structurally similar sentence, which we'll call an "instance."

[2]Many logic texts also refer to this as substitution.

Definition 2.1.8 (Instance of a Sentence).
Given a sentence, A, in L_V, there will be a set of atomic sentences that occur in A. An instance of A is a sentence (in L_V) given by substituting each occurrence of an atomic sentence, $\mathbb{P}_i$ (in A), with a corresponding sentence, D_i (in L_V), according to a function

$$s : \textsc{Props} \to L_V$$
$$s : \mathbb{P}_i \mapsto D_i.$$

Let's go over some examples before moving on.

Example 2.1.9.
Consider the sentence, $A = (\mathbb{P}_1 \wedge \mathbb{P}_2) \to \mathbb{P}_2$. The following are instances of A,

$$(\mathbb{P}_1 \wedge \mathbb{P}_1) \to \mathbb{P}_1, \qquad ((\Box \mathbb{P}_4 \vee \neg\mathbb{P}_2) \wedge (\mathbb{P}_2 \prec \mathbb{P}_3)) \to (\mathbb{P}_2 \prec \mathbb{P}_3).$$

In the first case, both $\mathbb{P}_1$ and $\mathbb{P}_2$ are substituted with $D_1 = D_2 = \mathbb{P}_1$. In the second case, $\mathbb{P}_1$ is substituted with $D_1 = (\Box \mathbb{P}_4 \vee \neg\mathbb{P}_2)$ and $\mathbb{P}_2$ is substituted with $D_2 = (\mathbb{P}_2 \prec \mathbb{P}_3)$.

The following are **not instances** of A,

$$(\mathbb{P}_1 \wedge \mathbb{P}_2) \to \mathbb{P}_3, \qquad (\mathbb{P}_2 \wedge \mathbb{P}_1) \to \mathbb{P}_2.$$

The first one fails because one occurrence of $\mathbb{P}_2$ is substituted with $\mathbb{P}_2$ and another is substituted with $\mathbb{P}_3$ (i.e., it is not clear what D_2 is supposed to be). The second one fails for a similar reason, it is unclear what the sentence D_2 is.

The vast majority of the time we will use the following notation to refer to instances of sentences.

Remark 2.1.10. Given a sentence, A, and atomic sentences, $p_1, \ldots, p_n \in \textsc{Props}$, we'll sometimes write

$$A[p_1/D_1, p_2/D_2, \ldots, p_n/D_n],$$

to refer to an instance of A obtained by substituting according to a function of the form:

$$s : \textsc{Props} \to L_V$$
$$s : p_i \mapsto D_i.$$

Classical Sentences

We will at times want to refer to classical propositional sentences in L_V, meaning sentences which only contain the classical connectives. To make this precise, we will define the sublanguage of L_V consisting of strings that only contain classical connectives.

Definition 2.1.11 (Language, L_0, for Classical Propositional Logic)**.**
Given the alphabet $(\mathscr{A} - \{ \preccurlyeq \})$, define the language L_0 so that

1. Props $\subseteq L_0$
2. a) $\bot \in L_0$
   b) If $A, B \in L_0$ then $(A \to B) \in L_0$.

No additional strings are included in L_0. We also adopt the additional *defined connectives* $\neg, \top, \wedge, \vee, \leftarrow, \leftrightarrow$ according to definition 2.1.6 but we adapt no other connectives.

The language L_0 is both a subset of L_V, by construction, and it is the language for classical propositional logic. When we refer to a **classical propositional sentence in L_V** we are referring specifically to a sentence in the sublanguage L_0. Moreover, a **tautology of classical propositional logic** will refer to a classical propositional sentence in L_V that is also a tautology in classical propositional logic.

2.2 Sphere Frames

The language we just described only gives us a collection of sentences formed out of an alphabet of symbols. It doesn't tell us anything about what those symbols mean or give us any notion of truth. Throughout the rest of this chapter we will go over how to assign a *semantics* to our language that does precisely this. We will define mathematical objects called *models* that assign truth values to our sentences and in doing so our definition of truth will characterize the behavior of our symbols. For instance, our $\to$ connective will behave, with respect to our truth values, as we are familiar in classical logic.

We will see that each model will contain a set of points, which we will call worlds, and we will be able to say whether or not a sentence is true at a specific world in a model. It's helpful to give a preview of the notation we'll be defining later. To say that a sentence is true specifically in a world w, within a model M, we will use the notation $M, w \vDash A$, this may be read as "M, w models A." Similarly, to say that the sentence is true at every world throughout the entire model, we will write $M \vDash A$.

In addition to worlds, our models will also come equipped with a structure on our set of worlds called a "system of spheres." This structure will describe a relationship between worlds. For some of our connectives, it will be the case that their behavior, with respect to truth values, will depend on this structure.

We are also interested in talking about sentences which are true across a family of models and for that purpose we will define a more general mathematical object called a *frame* upon which a family of models can be generated. Similarly, we'll also define truth over a family of frames, this will be very useful for later chapters.

We will begin by defining a sphere frame.

Definition 2.2.1 (Sphere frame)**.**
A **sphere frame** is a pair $F = (W, \mathcal{O})$ where:

1. W is a set, called a *set of worlds*,
2. $\mathcal{O} : W \to \mathscr{P}(\mathscr{P}(W))$ is a function called a *system of spheres*.

   For each $w \in W$, $\mathcal{O}(w)$ is called a *set of spheres around* w and it satisfies the following properties:

   a) $\mathcal{O}(w)$ is totally ordered by $\subseteq$, and
   b) $\mathcal{O}(w)$ is closed under arbitrary[3] union and non-empty[4] intersection (possibly infinite).

[3] The empty union is the empty set.

[4] Empty intersection is not allowed because it forces us to include the set W as a sphere around any world. This would greatly restrict the sort of models we could describe.

At times we will simply refer to these as **frames**.

This definition is extremely general and it's worth making some remarks about this.

Remark 2.2.2. It is *not necessarily the case* that:

1. knowing what $\mathcal{O}(w)$ looks like will tell us *anything* about what $\mathcal{O}(v)$ should look like – both could look wildly different.
2. $\{ w \} \in \mathcal{O}(w)$.
3. $W \in \mathcal{O}(w)$.
4. $w \in \bigcup \mathcal{O}(w)$.
5. $W = \bigcup \mathcal{O}(w)$, i.e., there could exist some $v \in W$ such that $v \notin \bigcup \mathcal{O}(w)$.

It is helpful to consider some examples to see that closure under union and intersection is not implied by the ordering of spheres. These examples are somewhat pathological but they serve to clarify the generality of the definition.

Example 2.2.3.
Let $W = \mathbb{R}_{>0}$, where $\mathbb{R}_{>0} = \{ r \in \mathbb{R} \mid r > 0 \}$. Then consider the frame $F = (\mathbb{R}_{>0}, \mathcal{O})$ where, for each $r \in \mathbb{R}_{>0}$:

$$\mathcal{O}(r) = \{ (r, s) \mid s \in \mathbb{R}, r < s \} .$$

There are a few observations we can make about this frame. Since our set of worlds is $\mathbb{R}$ then we may visualize it as a number line, with spheres as open intervals. Consider the case $3 \in W$, in this case $\mathcal{O}(3)$ contains every open interval of the form $(3, s)$ where $s \in \mathbb{R}$, refer to fig. 2.1. Taking the union over

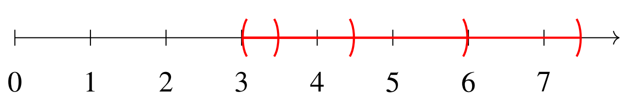


Figure 2.1: Some of the spheres in $\mathcal{O}(3)$.

all spheres gives us $\bigcup \mathcal{O}(3) = (3, \infty)$. Similarly, taking the intersection over all spheres gives us the interval $\bigcap \mathcal{O}(3) = \varnothing$. Some further observations are that every non-empty sphere has infinite cardinality, $\{ 3 \} \notin \mathcal{O}(3)$, $W \notin \mathcal{O}(3)$, and there is no sphere in $\mathcal{O}(3)$ containing 3 (i.e., 3 is, in a sense, inaccessible from 3).

Furthermore, we can note that there is no *smallest sphere* in $\mathcal{O}(3)$ that intersects the set $(4, \infty)$. To see why, consider that the set of all spheres in $\mathcal{O}(3)$ that intersect $(4, \infty)$ is $\{ (3, s) \mid s \in \mathbb{R} \text{ and } 4 < s \}$, see fig. 2.2, but taking the intersection yields the set $(3, 4]$ which no longer intersects with $(4, \infty)$.

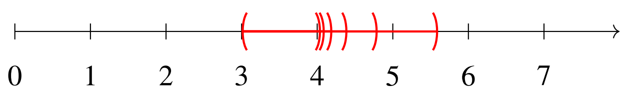


Figure 2.2: Some of the spheres in $\mathcal{O}(3)$ that intersect the set $(4, \infty)$.

This second pathological example is similar to the first one but it exhibits some different, potentially counterintuitive, properties.

Example 2.2.4.
Let $W = \mathbb{Z}$. Then consider the frame $F = (\mathbb{Z}, \mathcal{O})$ where, for each $n \in \mathbb{Z}$:

$$\mathcal{O}(n) = \{ (-\infty, m] \cap \mathbb{Z} \mid m \in \mathbb{Z}, m \leq n \} .$$

Like our previous example we can visualize this on a number line. Consider the case $3 \in W$, in this case, for each integer $m \leq 3$, $\mathcal{O}(3)$ contains the set $\{ m, m-1, m-2, \dots \}$, refer to fig. 2.3. Taking the union over all sets gives us

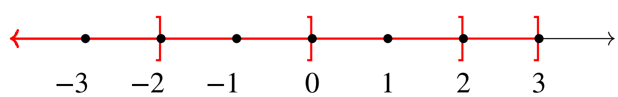


Figure 2.3: Some of the spheres in $\mathcal{O}(3)$.

$(-\infty, 3] \cap \mathbb{Z}$. Taking the intersection of all sets gives us the empty set. There is only one sphere in $\mathcal{O}(3)$ that contains 3 itself and that is the outermost sphere, $(-\infty, 3] \cap \mathbb{Z}$. There are no spheres in $\mathcal{O}(3)$ containing 4, so 4 is, in a sense, inaccessible from 3.

2.3 Sphere Models

Next we would like to define sphere models, but first some notation.

Definition 2.3.1 (Non-empty Intersection)**.**
For any sets, A and B, define the following meta-logical operation,

$$(A \bullet B) \text{ iff } (A \cap B) \neq \varnothing,$$

and read it as as "*A intersects B*." Similarly, define,

$$(A \not\bullet B) \text{ iff } \neg(A \bullet B),$$

and read it as "*A does not intersect B*."

Now we give the definition of a sphere model.

Definition 2.3.2 (Sphere Model)**.**
A **sphere model** over a frame, $F = (W, \mathcal{O})$, is a triple, $M = (W, \mathcal{O}, v)$, where v is a function, called a *valuation*,[5] such that:

$$v : \{ \mathbb{P}_n \mid n \in \mathbb{N} \} \to \mathcal{P}(W).$$

At times we will simply refer to these as **models**.

We may think of a valuation as assigning, to each atomic sentence, a set of worlds where that atomic sentence is true. We will formally define truth later, but first we will generalize this notion of valuation to all sentences (not just atomic sentences).

Definition 2.3.3 (Generalized Valuation)**.**
Given a sphere model, $M = (W, \mathcal{O}, v)$, and a sentence, A, we will define a generalized valuation function,

$$[\![\cdot]\!] : L_V \to \mathcal{P}(W),$$

recursively as follows:

[5]The term valuation is commonly used in logic for a function that assigns truth values to sentences. It is unrelated to the notion of a valuation in algebra.

1. $[\![\mathbb{P}_n]\!] = v(\mathbb{P}_n)$,

2. $[\![\bot]\!] = \varnothing$,

3. $[\![A \to B]\!] = (W - [\![A]\!]) \cup [\![B]\!]$,

4. $[\![A \preccurlyeq B]\!] = \{\, w \in W \mid \forall S \in \mathcal{O}(w), (([\![B]\!] \bullet S) \to ([\![A]\!] \bullet S))\,\}$.[6]

Remark 2.3.4. Lewis defines sphere models a little differently (Lewis, 2001, Section 6.1). Instead of using a valuation he directly uses a generalized valuation. Additionally, he includes all of the connectives as part of the definition whereas we split them off into the following theorem. We've changed things in order to be consistent with modern logic texts.

Since there are several defined connectives that we are interested in, it is helpful to prove a number of theorems characterizing their behavior before we move onto defining truth.

Theorem 2.3.5.
In a sphere model, $M = (W, \mathcal{O}, v)$, the following also hold for any sentences A and B:

1. $[\![\neg A]\!] = W - [\![A]\!]$

2. $[\![\neg\neg A]\!] = [\![A]\!]$

3. $[\![A \to B]\!] = [\![\neg A]\!] \cup [\![B]\!]$

4. $[\![\top]\!] = W$

5. $[\![A \wedge B]\!] = [\![A]\!] \cap [\![B]\!]$

6. $[\![A \vee B]\!] = [\![A]\!] \cup [\![B]\!]$

7. $[\![A \leftrightarrow B]\!] = ([\![A]\!] \cap [\![B]\!]) \cup ([\![\neg A]\!] \cap [\![\neg B]\!])$

8. $[\![A \prec B]\!] = \{\, w \in W \mid \exists S \in \mathcal{O}(w), ([\![A]\!] \bullet S \wedge [\![B]\!] \not\bullet S)\,\}$

[6] The statement within the set-builder notation is not a sentence in counterfactual logic. It is an ordinary classical sentence in the meta-theory.

9. $[\![A \approx B]\!] = \{ \ w \in W \mid \forall S \in \mathcal{O}(w), ([\![B]\!] \bullet S \leftrightarrow [\![A]\!] \bullet S) \ \}$

10. $[\![\Diamond A]\!] = \left\{ \ w \in W \ \middle| \ [\![A]\!] \bullet \bigcup \mathcal{O}(w) \ \right\}$

11. $[\![\Box A]\!] = \left\{ \ w \in W \ \middle| \ \bigcup \mathcal{O}(w) \subseteq [\![A]\!] \ \right\}$

12. $[\![\Diamond\!\!\cdot A]\!] = \{ \ w \in W \mid \forall S \in \mathcal{O}(w), S \neq \varnothing \rightarrow [\![A]\!] \bullet S \ \}$

13. $[\![\boxdot A]\!] = \{ \ w \in W \mid \exists S \in \mathcal{O}(w), S \neq \varnothing \wedge S \subseteq [\![A]\!] \ \}$

14. $$[\![A \mathrel{\Box\!\!\Rightarrow} B]\!] = \left\{ \ w \in W \ \middle| \ \begin{array}{l} \exists S \in \mathcal{O}(w), \\ \quad (([\![A]\!] \bullet S) \wedge (([\![A]\!] \cap S) \subseteq [\![B]\!])) \end{array} \right\}$$

15. $$[\![A \mathrel{\Diamond\!\!\Rightarrow} B]\!] = \left\{ \ w \in W \ \middle| \ \begin{array}{l} \forall S \in \mathcal{O}(w), \\ \quad ([\![A]\!] \bullet S \rightarrow ([\![A]\!] \cap [\![B]\!]) \bullet S) \end{array} \right\}$$

16. $$[\![A \mathrel{\Box\!\!\rightarrow} B]\!] = \left\{ \ w \in W \ \middle| \ \begin{array}{l} [\![A]\!] \bullet \bigcup \mathcal{O}(w) \rightarrow \exists S \in \mathcal{O}(w), \\ \quad (([\![A]\!] \bullet S) \wedge (([\![A]\!] \cap S) \subseteq [\![B]\!])) \end{array} \right\}$$

17. $$[\![A \mathrel{\Diamond\!\!\rightarrow} B]\!] = \left\{ \ w \in W \ \middle| \ \begin{array}{l} [\![A]\!] \bullet \bigcup \mathcal{O}(w) \wedge \forall S \in \mathcal{O}(w), \\ \quad ([\![A]\!] \bullet S \rightarrow ([\![A]\!] \cap [\![B]\!]) \bullet S) \end{array} \right\}$$

Proof. Let $M = (W, \mathcal{O}, v)$ be a sphere model and let A and B be arbitrary sentences. Let $w \in W$ be arbitrary.[7]

In order to prevent clutter, we will only present proofs of two of these statements here. A full presentation of the remaining proofs may be found in the Additional Proofs chapter theorem 8.1.2.

[7]Since the domain of generalized valuation is subsets of W, then this assumption doesn't weaken our proofs. It will, however, streamline our proofs by allowing us to write $w \notin [\![A]\!]$ instead of $w \in (W - [\![A]\!])$ and letting us use statements like $w \in \{v \in W \mid \phi(v)\}$ iff $\phi(w)$.

1.

$$
\begin{aligned}
&[\![\neg A]\!] \\
&\quad = [\![A \to \bot]\!] \\
&\qquad \text{(by definition of } \neg A\text{, definition 2.1.6)} \\
&\quad = (W - [\![A]\!]) \cup [\![\bot]\!] \\
&\qquad \text{(by definition of } [\![A \to B]\!]\text{, definition 2.3.3(item 3))} \\
&\quad = (W - [\![A]\!]) \cup \varnothing \\
&\qquad \text{(by definition of } [\![\bot]\!]\text{, definition 2.3.3(item 2))} \\
&\quad = (W - [\![A]\!]).
\end{aligned}
$$

Effectively, this means that $[\![\neg A]\!]$ is the complement of $[\![A]\!]$ with respect to W.

11.

$$
\begin{aligned}
&w \in [\![\Box A]\!] \\
&\quad \text{iff } w \in [\![\bot \preccurlyeq \neg A]\!] \\
&\qquad \text{(by definition of } \Box A\text{, definition 2.1.6)} \\
&\quad \text{iff } w \in \{\, v \in W \mid \forall S \in \mathcal{O}(v), ([\![\neg A]\!] \bullet S \to [\![\bot]\!] \bullet S) \,\} \\
&\qquad \text{(by definition of } [\![A \preccurlyeq B]\!]\text{, definition 2.3.3(item 4))} \\
&\quad \text{iff } \forall S \in \mathcal{O}(w), ([\![\neg A]\!] \bullet S \to [\![\bot]\!] \bullet S) \\
&\quad \text{iff } \forall S \in \mathcal{O}(w), ([\![\neg A]\!] \bullet S \to \varnothing \bullet S) \\
&\qquad \text{(by definition of } [\![\bot]\!]\text{, definition 2.3.3(item 2))} \\
&\quad \text{iff } \forall S \in \mathcal{O}(w), ([\![\neg A]\!] \not\bullet S) \\
&\quad \text{iff } \forall S \in \mathcal{O}(w), (S \subseteq [\![A]\!]) \\
&\qquad \text{(by lemma 8.1.1)} \\
&\quad \text{iff } \bigcup \mathcal{O}(w) \subseteq [\![A]\!] \\
&\quad \text{iff } w \in \left\{\, v \in W \;\middle|\; \bigcup \mathcal{O}(v) \subseteq [\![A]\!] \,\right\}.
\end{aligned}
$$

□

Lemma 2.3.6.
Given a sphere model, $M = (W, \mathcal{O}, v)$*,*

$$[\![A]\!] \subseteq [\![B]\!] \textit{ iff } [\![A \to B]\!] = W.$$

Proof. We prove each direction separately.

For the first direction, suppose $[\![A]\!] \subseteq [\![B]\!]$.

$w \in [\![A \to B]\!]$

iff $w \in ([\![\neg A]\!] \cup [\![B]\!])$

(by theorem of $[\![A \to B]\!]$, theorem 2.3.5(item 3))

only if $w \in ([\![\neg A]\!] \cup ([\![A]\!] \cup [\![B]\!]))$

(since $[\![A]\!] \subseteq [\![B]\!]$ implies that $[\![B]\!] = [\![A]\!] \cup [\![B]\!]$)

iff $w \in (W \cup [\![B]\!])$

(by theorem of $[\![\neg A]\!]$, theorem 2.3.5(item 1))

iff $w \in W$.

For the second direction, suppose $[\![A \to B]\!] = W$. Suppose $w \in [\![A]\!]$. By assumption, we have $w \in [\![A \to B]\!]$. So, by theorem 2.3.5(item 3), $w \in ([\![\neg A]\!] \cup [\![B]\!])$. However, since $w \in [\![A]\!]$ then, by theorem 2.3.5(item 3), $w \notin [\![\neg A]\!]$, meaning that we must have $w \in [\![B]\!]$. Since we assumed $w \in [\![A]\!]$ was arbitrary then we have $[\![A]\!] \subseteq [\![B]\!]$. $\square$

In *Counterfactuals* (Lewis, 2001), Lewis defines $\Box A$ as $\bot \approx \neg A$ and $\Diamond A$ as $A \approx \top$. Since we define them differently in definition 2.1.6, we will take the time to show that both pairs of definitions are equivalent in our semantics.

Theorem 2.3.7.
Given a sphere model, $M = (W, \mathcal{O}, v)$*. The following are true for all sentences* A*.*

1. $[\![\neg A \preccurlyeq \bot]\!] = W$
2. $[\![\top \preccurlyeq A]\!] = W$
3. $[\![\bot \approx \neg A]\!] = [\![\Box A]\!]$

4. $[\![A \approx \top]\!] = [\![\Diamond A]\!]$

Proof. Let $M = (W, \mathcal{O}, v)$ be a sphere model and let A be an arbitrary sentence. We will prove each sentence separately.

1.

$$\begin{aligned}
&[\![\neg A \preccurlyeq \bot]\!] \\
&\quad = \{\, v \in W \mid \forall S \in \mathcal{O}(v), (([\![\bot]\!] \bullet S) \rightarrow ([\![\neg A]\!] \bullet S)) \,\} \\
&\qquad \text{(by definition of } [\![A \preccurlyeq B]\!]\text{, definition 2.3.3(item 4))} \\
&\quad = \{\, v \in W \mid \forall S \in \mathcal{O}(v), ((\varnothing \bullet S) \rightarrow ([\![\neg A]\!] \bullet S)) \,\} \\
&\quad = W \\
&\qquad \text{(since } \varnothing \not\bullet S\text{).}
\end{aligned}$$

2.

$$\begin{aligned}
&[\![\top \preccurlyeq A]\!] \\
&\quad = \{\, v \in W \mid \forall S \in \mathcal{O}(v), (([\![A]\!] \bullet S) \rightarrow ([\![\top]\!] \bullet S)) \,\} \\
&\qquad \text{(by definition of } [\![A \preccurlyeq B]\!]\text{, definition 2.3.3(item 4)).}
\end{aligned}$$

Here we have two cases. If $S = \varnothing$ then $[\![A]\!] \bullet S$ is false, and if $S \neq \varnothing$ then $W \bullet S$. In either case, $(([\![A]\!] \bullet S) \rightarrow ([\![\top]\!] \bullet S))$ is true. So

$$\{\, v \in W \mid \forall S \in \mathcal{O}(v), (([\![A]\!] \bullet S) \rightarrow ([\![\top]\!] \bullet S)) \,\} = W.$$

3.

$$
\begin{aligned}
&[\![\bot \approx \neg A]\!] \\
&\quad = [\![(\bot \preccurlyeq \neg A) \wedge (\neg A \preccurlyeq \bot)]\!] \\
&\qquad \text{(by definition of } A \approx B\text{, definition 2.1.6)} \\
&\quad = [\![\bot \preccurlyeq \neg A]\!] \cap [\![\neg A \preccurlyeq \bot]\!] \\
&\qquad \text{(by theorem of } [\![A \wedge B]\!]\text{, theorem 2.3.5(item 5))} \\
&\quad = [\![\bot \preccurlyeq \neg A]\!] \cap W \\
&\qquad \text{(by theorem of } [\![\neg A \preccurlyeq \bot]\!]\text{, theorem 2.3.7(item 1))} \\
&\quad = [\![\bot \preccurlyeq \neg A]\!] \\
&\quad = [\![\Box A]\!] \\
&\qquad \text{(by definition of } \Box A\text{, definition 2.1.6).}
\end{aligned}
$$

4.

$$
\begin{aligned}
&[\![A \approx \top]\!] \\
&\quad = [\![(A \preccurlyeq \top) \wedge (\top \preccurlyeq A)]\!] \\
&\qquad \text{(by definition of } A \approx B\text{, definition 2.1.6)} \\
&\quad = [\![A \preccurlyeq \top]\!] \cap [\![\top \preccurlyeq A]\!] \\
&\qquad \text{(by theorem of } [\![A \wedge B]\!]\text{, theorem 2.3.5(item 5))} \\
&\quad = [\![A \preccurlyeq \top]\!] \cap W \\
&\qquad \text{(by theorem of } [\![\top \preccurlyeq A]\!]\text{, theorem 2.3.7(item 1))} \\
&\quad = [\![A \preccurlyeq \top]\!] \\
&\quad = [\![\diamond A]\!] \\
&\qquad \text{(by definition of } \diamond A\text{, definition 2.1.6).}
\end{aligned}
$$

□

2.4 Truth and Validity

We now describe truth and validity for worlds, models, sphere frames, and classes of sphere frames.

The following definition lets us generalize valuation even further over sets of sentences. This will help streamline the notation going forward.

Definition 2.4.1.
Given a model, $M = (W, \mathcal{O}, v)$, and a set of sentences Γ, define $[\![\Gamma]\!]$, such that $\forall w \in W$,

$$w \in [\![\Gamma]\!] \quad \text{iff} \quad \forall A \in \Gamma, w \in [\![A]\!].$$

This is the set of worlds where every sentence in Γ is true.

Generalized valuation only lets us refer to individual sentences, but this lets us refer to entire sets of sentences, possibly with an infinite cardinality.

Theorem 2.4.2.
Given a model, $M = (W, \mathcal{O}, v)$ and a set of sentences Γ, the following are true:

1. *If $\Gamma = \varnothing$ then $[\![\Gamma]\!] = W$.*
2. *If $\Gamma = \{ A \}$ then $[\![\Gamma]\!] = [\![A]\!]$, so we may omit curly braces when Γ is a singleton set.*
3. *If $\Gamma \neq \varnothing$ then*
$$[\![\Gamma]\!] = \bigcap_{A \in \Gamma} [\![A]\!].$$

Proof. The proofs follow trivially from the definition. □

Definition 2.4.3 (Truth and Validity)**.**
Given a set of sentences, Γ, we define the following:

1. Given a model, $M = (W, \mathcal{O}, v)$, and a world, $w \in W$,
$$M, w \vDash \Gamma \quad \text{iff} \quad w \in [\![\Gamma]\!].$$
Read: Γ **is true at the world,** w**, in the model,** M.
2. Given a model, $M = (W, \mathcal{O}, v)$,
$$M \vDash \Gamma \quad \text{iff} \quad \forall w \in W, M, w \vDash \Gamma.$$
Read: Γ **is valid over the model,** M.

3. Given a sphere frame, F,

$$F \vDash \Gamma \quad \text{iff} \quad \forall M \text{ over } F, M \vDash \Gamma.$$

Read: Γ **is valid over the sphere frame, F**.

4. Given a class of sphere frames, $\mathscr{F}$,

$$\vDash_{\mathscr{F}} \Gamma \quad \text{iff} \quad \forall F \in \mathscr{F}, F \vDash \Gamma.$$

Read: Γ **is valid over the class of sphere frames,** $\mathscr{F}$.

5.

$$\vDash \Gamma \quad \text{iff} \quad \vDash_{\mathscr{F}} \Gamma \text{ where } \mathscr{F} \text{ is the class of all sphere frames.}$$

Read: Γ **is valid**.

For each of these definitions, if Γ contains a single sentence, then we may omit the curly braces. For instance, we would write $M, w \vDash A$ instead of $M, w \vDash \{ A \}$.

Theorem 2.4.4.
Given a set of sentences, Γ, and a model $M = (W, \mathcal{O}, v)$,

$$M \vDash \Gamma \textit{ iff } [\![\Gamma]\!] = W.$$

Proof. The proof follows trivially from the definitions. □

We now prove that if a sentence is valid then all instances of that sentence are valid.

Theorem 2.4.5.
Given a frame F and a sentence A

$$\textit{if } F \vDash A, \textit{ then for every instance } A' \textit{ of } A, F \vDash A'.$$

Proof. Let F be a frame and A be a sentence. Suppose that $F \vDash A$.

Let A' be an arbitrary instance of A. Write:

$$A' = A[p_1/D_1, \dots, p_n/D_n]$$

where $p_1, \dots, p_n \in$ Props and $D_1, \dots, D_n$ are sentences. We wish to show that $F \vDash A[p_1/D_1, \dots, p_n/D_n]$.

Let $M = (W, \mathcal{O}, v)$ be an arbitrary model over F. Next, define a new valuation v' such that for $p_1, \dots, p_n \in$ Props, we have $v'(p_i) = [\![D_i]\!]$ (where $[\![D_i]\!]$ is defined in terms of our original valuation v) and for any other $q \in$ Props we have $v'(q) = [\![q]\!]$. With this valuation we can obtain another sphere model over F, $M' = (W, \mathcal{O}, v')$. Denote the generalized valuation for M' with $[\![\cdot]\!]'$.

Next, we proceed by induction on the complexity of A using the definition of generalized valuation (2.3.3), to prove that that $[\![A']\!] = [\![A]\!]'$.

Base case:

1. If $A := \bot$, then

$$[\![A']\!] = [\![A[p_1/D_1, \dots, p_n/D_n]]\!] = [\![\bot]\!] = [\![\bot]\!]' = [\![A]\!]'.$$

2. If $A := p_i$, then

$$[\![A']\!] = [\![A[p_1/D_1, \dots, p_n/D_n]]\!] = [\![D_i]\!] = [\![p_i]\!]' = [\![A]\!]'.$$

3. If $A := q$ (where $q \neq p_i$), then

$$[\![A']\!] = [\![A[p_1/D_1, \dots, p_n/D_n]]\!] = [\![D_i]\!] = [\![q]\!]' = [\![A]\!]'.$$

Induction step: Suppose that

$$[\![B[p_1/D_1, \dots, p_n/D_n]]\!] = [\![B]\!]' \text{ and } [\![C[p_1/D_1, \dots, p_n/D_n]]\!] = [\![C]\!]'.$$

1. If $A := B \to C$, then

$$\begin{aligned}
[\![A']\!] &= [\![A[p_1/D_1, \dots, p_n/D_n]]\!] \\
&= [\![(B \to C)[p_1/D_1, \dots, p_n/D_n]]\!] \\
&= [\![(B[p_1/D_1, \dots, p_n/D_n] \to C[p_1/D_1, \dots, p_n/D_n])]\!] \\
&= (W - [\![B[p_1/D_1, \dots, p_n/D_n]]\!]) \cup [\![C[p_1/D_1, \dots, p_n/D_n]]\!] \\
&= (W - [\![B]\!]') \cup [\![C]\!]' \\
&= [\![(B \to C)]\!]' \\
&= [\![A]\!]'.
\end{aligned}$$

2. If $A := B \preccurlyeq C$, then

$$\begin{aligned}
[\![A']\!] &= [\![A[p_1/D_1, \dots, p_n/D_n]]\!] \\
&= [\![(B \preccurlyeq C)[p_1/D_1, \dots, p_n/D_n]]\!] \\
&= [\![(B[p_1/D_1, \dots, p_n/D_n] \preccurlyeq C[p_1/D_1, \dots, p_n/D_n])]\!] \\
&= \left\{ w \in W \;\middle|\; \begin{array}{l} \forall S \in \mathcal{O}(w), \\ ([\![C[p_1/D_1, \dots, p_n/D_n]]\!] \bullet S \to \\ \quad [\![B[p_1/D_1, \dots, p_n/D_n]]\!] \bullet S) \end{array} \right\} \\
&= \left\{ w \in W \;\middle|\; \forall S \in \mathcal{O}(w), ([\![C]\!]' \bullet S \to [\![B]\!]' \bullet S) \right\} \\
&= [\![(B \preccurlyeq C)]\!]' \\
&= [\![A]\!]'.
\end{aligned}$$

This completes our induction showing that $[\![A']\!] = [\![A]\!]'$.

Finally, since A is valid over F (and hence over M'), then $W = [\![A]\!]' = [\![A']\!]$. Therefore $[\![A']\!]$ is valid over M, and since M was an arbitrary model over F it follows that $[\![A']\!]$ is valid over F. Moreover, since $[\![A']\!]$ was an arbitrary substitution instance of A then the result holds for all substitution instances. □

Next, we show that all classical tautologies are valid as well as all instances of classical tautologies.

Lemma 2.4.6.
Classical tautologies are valid over every sphere model.

Proof. Consider an arbitrary sphere model and note that classical tautologies are valid at each world. Since they are valid at every world then they are valid over the entire sphere model, and since this was an arbitrary sphere model the result holds over all sphere models. □

Theorem 2.4.7.
All instances of classical tautologies are valid in every sphere model.

Proof. Let F be an arbitrary frame. Since all classical tautologies are valid over F then all instances are valid over F. Since F was arbitrary then all instances of classical tautologies are valid over every sphere model. □

2.5 Properties of Frames and Models

Lewis (2001) uses a variety of properties to describe the semantics for a family of 26 logics. We distinguish between properties of frames and properties of models.[8]

Definition 2.5.1 (Properties of Frames)**.**
The properties in table table 2.2[9,10] are all properties of frames. That means that if a frame satisfies one of those properties, then every model over that frame will also satisfy those properties. Later, we will be interested in classes of frames, where every frame in the class satisfies the property.

We may refer to a property by either its long name or short name. For instance, one could say that a frame is centered instead of saying that said frame satisfies the property **C**.

Definition 2.5.2 (Properties of Models)**.**
The properties in table table 2.2 are properties of models because they reference the generalized valuation of sentences. In general, one cannot specify a class of frames over which these properties are true. Many of our major results later on that refer to frames can be restated in terms of classes of models. I have avoided doing this, however, because in our semantics frames are, in some sense, more logically meaningful than classes of models.

For convenience, we may refer to a model, frame, or class of frames that satisfies a sequence of properties, $\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_n$, as a $\mathbf{VP}_1\mathbf{P}_2 \cdots \mathbf{P}_n$ model, frame, or class of frames. For instance, a **VCSU** model will refer to a model that satisfies properties, **C**, **S**, and **U**.

I do not want to give the reader the impression that the Limit Assumption and Stalnaker's Assumption aren't useful or interesting. On the contrary, they have a very rich history and allow one to reformulate our semantics in interesting ways.

[8] Lewis doesn't define frames. He only refers to classes of models, so for him there is no distinction.

[9] Lewis does not provide the properties **4** and **5** but they can be obtained by splitting **U-** in half.

[10] Lewis defines these properties in prose our definitions formalize Lewis' descriptions.

Normal	**N**	$\forall w \in W$, $\left(\bigcup \mathcal{O}(w)\right) \neq \varnothing$
Reflexive	**T**	$\forall w \in W$, $w \in \left(\bigcup \mathcal{O}(w)\right)$
Weakly Centered	**W**	$\forall w \in W$, $w \in \left(\bigcap (\mathcal{O}(w) - \{ \varnothing \})\right)$ and $\mathcal{O}(w) \neq \{ \varnothing \}$
Centered	**C**	$\forall w \in W$, $\{ w \} \in \mathcal{O}(w)$
Locally Uniformity	**U-**	$\forall w \in W, \forall v \in \left(\bigcup \mathcal{O}(w)\right)$, $\left(\bigcup \mathcal{O}(v)\right) = \left(\bigcup \mathcal{O}(w)\right)$
Uniform	**U**	$\forall w, v \in W$, $\left(\bigcup \mathcal{O}(v)\right) = \left(\bigcup \mathcal{O}(w)\right)$
Locally Absolute	**A-**	$\forall w \in W, \forall v \in \left(\bigcup \mathcal{O}(w)\right)$, $\mathcal{O}(v) = \mathcal{O}(w)$
Absolute	**A**	$\forall w, v \in W$, $\mathcal{O}(v) = \mathcal{O}(w)$
Universal	**UT**	$\forall w \in W$, $\left(\bigcup \mathcal{O}(w)\right) = W$
Weakly Trivial	**WA**	$\forall w \in W$, $\mathcal{O}(w) = \{ W \}$ or $\mathcal{O}(w) = \{ W, \varnothing \}$
Trivial	**CA**	$\forall w \in W$, $W = \{ w \}$ and $\left(\bigcup \mathcal{O}(w)\right) = W$
Transitive	**4**	$\forall w \in W, \forall v \in \left(\bigcup \mathcal{O}(w)\right)$, $\left(\bigcup \mathcal{O}(v)\right) \subseteq \left(\bigcup \mathcal{O}(w)\right)$
Euclidean	**5**	$\forall w \in W, \forall v \in \left(\bigcup \mathcal{O}(w)\right)$, $\left(\bigcup \mathcal{O}(v)\right) \supseteq \left(\bigcup \mathcal{O}(w)\right)$

Table 2.2: Properties of frames.

Limit Assumption	**L**	$\forall A \in L_V, \forall w \in W$, if $\left(\bigcup \mathcal{O}(w)\right) \bullet [\![A]\!]$ then $\left(\bigcap \{ S \in \mathcal{O}(w) \mid S \bullet [\![A]\!] \}\right) \bullet [\![A]\!]$
Stalnaker's Assumption	**S**	$\forall A \in L_V, \forall w \in W$, if $\left(\bigcup \mathcal{O}(w)\right) \bullet [\![A]\!]$ then $\left|\left(\bigcap \{ S \in \mathcal{O}(w) \mid S \bullet [\![A]\!] \}\right) \bigcap [\![A]\!]\right| = 1$

Table 2.3: Properties of models.

Remark 2.5.3. Models that satisfy the Limit Assumption are exceptionally well behaved. In such models one is always allowed to talk about a *most similar set* *of worlds* where a given sentence is true. This makes it possible to define a "set-selection function", $f : L_v \times W \to \mathscr{P}(W)$, such that $f(A, w)$ is equal to the most similar worlds where A is true. Lewis discusses this in *Counterfactuals* (Lewis, 2001, Section 2.7) but this is also the basis of how Chellas (1975) approaches counterfactuals. Another thing one can do is to define, for each sentence A, a family of relations on $R_A \subseteq W \times W$ such that $(w, v) \in R_A$ iff v is one of the most similar worlds to w where A is true, this forms the basis of how Segerberg (1989) approaches counterfactual logic.

Models that satisfy Stalnaker's Assumption are even more well behaved. These are models where, from each world, if there exist any similar worlds where a given sentence is true, then there is exactly one "most similar" world where said sentence is true. This forms the basis of how Stalnaker (1968) approaches counterfactual logic. By including an additional "absurd world" where all sentences are true, Stalnaker is able to define a "selection function", $f : L_v \times W \to W$, such that $f(A, w) = v$ is interpreted to mean that v is the most similar world to w where A is true (and if v is the "absurd world" it is interpreted to mean that there are no similar worlds to w where A is true). Lewis discusses the relationship between his and Stalnaker's approaches in more detail within *Counterfactuals* (Lewis, 2001, Section 3.4).

2.6 Entailment from a Set of Sentences

We next introduce four notions describing the way that a set of sentences can entail a sentence. Usually only the first two are of interest but we include the last two so that we may have convenient notation for them later on but they are not usually studied in logic.

Definition 2.6.1 (Entailment)**.**
Given a sentence, A, a set of sentences, Γ, and a class of sphere frames, $\mathscr{F}$, we define two different kinds of entailment:

1. **Local entailment:** This version of entailment deals with **truth** at individual worlds. Γ *locally entails* A means that for any world in any model over any sphere frame in $\mathscr{F}$, if Γ is true at a world then A will also be

true at that world. Formally, we say that Γ **locally entails** A **over** $\mathscr{F}$ iff $\forall F = (W, \mathcal{O}) \in \mathscr{F}, \forall M$ over $F, \forall w \in W$,

$$\text{if } M, w \vDash \Gamma \text{ then } M, w \vDash A.$$

Write:

$$\Gamma \vDash^{\ell}_{\mathscr{F}} A.$$

Often we'll use this version of entailment when we want to talk about things **preserving truth**.

2. **Global Entailment:** This version of entailment deals with **validity** (over individual models). When Γ *globally entails* A it means that for any model over any sphere frame in $\mathscr{F}$, if Γ is valid over that model then A will also be valid over that model. Formally, we say that Γ **globally entails** A **over** $\mathscr{F}$ iff $\forall F = (W, \mathcal{O}) \in \mathscr{F}, \forall M$ over F,

$$\text{if } M \vDash \Gamma \text{ then } M \vDash A.$$

Write:

$$\Gamma \vDash^{g}_{\mathscr{F}} A.$$

Often we'll use this version of entailment when we want to talk about things **preserving validity**.

Theorem 2.6.2.
Given a sentence, A, a set of sentences, Γ, and a class of sphere frames, $\mathscr{F}$,

$$\textit{If } \Gamma \vDash^{\ell}_{\mathscr{F}} A \textit{ then } \Gamma \vDash^{g}_{\mathscr{F}} A,$$

Proof. First, expand the definitions of these statements and then note that the difference between is merely the scope of the universal quantifiers. Hence, our proof follows from an application of the following property of the universal quantifier:

$$\text{If } \forall x, (P(x) \to Q(x)) \text{ then } \forall x P(x) \to \forall x Q(x).$$ □

Theorem 2.6.3.
Given a sentence A and a class of sphere frames $\mathscr{F}$, each of the following statements are equivalent to $\vDash_{\mathscr{F}} A$:

1. $\varnothing \vDash^{\ell}_{\mathscr{F}} A$.

2. $\varnothing \vDash^{g}_{\mathscr{F}} A$.

Proof. This proof follows trivially. Again, expand the definitions and note that $M, w \vDash \varnothing$ is vacuously true. With that observation, both expressions collapses into the following statement

$$\forall F = (W, \mathscr{O}) \in \mathscr{F}, \forall M \text{ over } F, \forall w \in W, M, w \vDash A,$$

which is just the definitional expansion of $\vDash_{\mathscr{F}} A$. $\square$

It's important to note that theorem 2.6.2 only goes in one direction. It's particularly important to show that global entailment does not imply local entailment, so we include a proof of that here.

Theorem 2.6.4.
Given a sentence, A, a set of sentences, Γ, and a class of sphere frames, $\mathscr{F}$, $\Gamma \vDash^{g}_{\mathscr{F}} A$ does not necessarily imply $\Gamma \vDash^{\ell}_{\mathscr{F}} A$.

Proof. To show these we simply need to construct an example where global entailment is satisfied but local entailment fails. Consider the case where $A = \square \mathbb{P}_1$ and $\Gamma = \{ \mathbb{P}_1 \}$. Let $\mathscr{F}$ be the class of all frames.

This satisfies $\{ \mathbb{P}_1 \} \vDash^{g}_{\mathscr{F}} \square \mathbb{P}_1$. To see why, let $F \in \mathscr{F}$ be arbitrary and let M be an arbitrary model over F.

Then, suppose $M \vDash \{ \mathbb{P}_1 \}$. We wish to show that $M \vDash \square \mathbb{P}_1$. Since $M \vDash \{ \mathbb{P}_1 \}$ then $[\![\mathbb{P}_1]\!] = W$. Next, by theorem 2.3.5(item 11) we know that

$$[\![\square \mathbb{P}_1]\!] = \left\{ w \in W \;\middle|\; \bigcup \mathscr{O}(w) \subseteq [\![\mathbb{P}_1]\!] \right\}.$$

However, since $[\![\mathbb{P}_1]\!] = W$ and $\forall w \in W$ we have $\bigcup \mathscr{O}(w) \subseteq W$ then, it must follow that $[\![\square \mathbb{P}_1]\!] = W$. Therefore, $M \vDash \square \mathbb{P}_1$. Since M and F were arbitrary, then we have that $\{ \mathbb{P}_1 \} \vDash^{g}_{\mathscr{F}} \square \mathbb{P}_1$. In fact, we have that $\{ A \} \vDash^{g}_{\mathscr{F}} \square A$ for arbitrary A.

Next, we show that it is not the case that $\{ \mathbb{P}_1 \} \vDash^{\ell}_{\mathscr{F}} \square \mathbb{P}_1$. To do this we construct a counterexample. Let $M = (W, \mathscr{O}, v)$ where $W = \{ w, v \}$, $w, v \in$

$\mathcal{O}(w)$, and $v(\mathbb{P}_1) = \{ w \}$. Since $\mathcal{F}$ is the class of all frames, then M is a model over an $F \in \mathcal{F}$.

It follows that $M, v \not\models \mathbb{P}_1$, and $M, w \models \mathbb{P}_1$. Moreover, since $v \in \bigcup \mathcal{O}(w)$ but $v \notin [\![\mathbb{P}_1]\!]$, then by definition $w \notin [\![\Box \mathbb{P}_1]\!]$, meaning that $M, w \not\models \Box \mathbb{P}_1$. Therefore, since there exists a world, w, and a model M over a frame, $F \in \mathcal{F}$ such that $M, w \models \mathbb{P}_1$ but $M, w \not\models \Box \mathbb{P}_1$ it follows that $\{ \mathbb{P}_1 \} \not\models^{\ell}_{\mathcal{F}} \Box \mathbb{P}_1$, completing our proof. □

Previously we've introduced the notion of generalized valuation as a function that takes a sentence (or set of sentences) and gives us a set of worlds where that sentence (or set of sentences) is true. It will be helpful, for illustrative purposes to define different sorts of valuations that return models and frames where a sentence or set of sentences is valid.

Definition 2.6.5.
Given a set of sentences Γ, a class of sphere frames $\mathcal{F}$, and a sphere frame $F \in \mathcal{F}$, we define $(\!(\Gamma)\!)_F$, such that, $\forall M$ over F,

$$M \in (\!(\Gamma)\!)_F \quad \text{iff} \quad M \models \Gamma.$$

This is the set of models over F where Γ is valid (over the model).

The following theorem gives us a way to reason about the different kinds of entailment by referring to sets of worlds, models, and frames. We may even visualize the different kinds of entailments using Eulerian diagrams.

Theorem 2.6.6.
Given a sentence, A, a set of sentences Γ, and a class of sphere frames $\mathcal{F}$, the following are true:

1.
$$\Gamma \models^{\ell}_{\mathcal{F}} A \quad \textit{iff} \quad \forall F = (W, \mathcal{O}) \in \mathcal{F}, \forall M \textit{ over } F,\ [\![\Gamma]\!] \subseteq [\![A]\!].$$

In other words, Γ locally entails A over $\mathcal{F}$ iff for every model M, over an $F \in \mathcal{F}$, the set of worlds where Γ is true is a subset of the set of worlds where A is true

2.

$$\Gamma \vDash^{g}_{\mathscr{F}} A \quad \textit{iff} \quad \forall F \in \mathscr{F}, (\!|\Gamma|\!)_F \subseteq (\!|A|\!)_F.$$

In other words, Γ globally entails A over $\mathscr{F}$ iff the models over frames in $\mathscr{F}$ where Γ is valid (over the model) are a subset of models over frames in $\mathscr{F}$ where A is valid (over the model).

Proof. We prove each direction for each statement.

1.

$\Gamma \vDash^{\ell}_{\mathscr{F}} A$

iff $\forall F = (W, \mathscr{O}) \in \mathscr{F}, \forall M$ over F,

$\forall w \in W,$ [if $M, w \vDash \Gamma$ then $M, w \vDash A$]

(by definition of $\Gamma \vDash^{\ell}_{\mathscr{F}} A$, definition 2.6.1(item 1))

iff $\forall F = (W, \mathscr{O}) \in \mathscr{F}, \forall M$ over F,

$\forall w \in W,$ [if $w \in [\![\Gamma]\!]$ then $w \in [\![A]\!]$]

(by definition of $M, w \vDash \Gamma$, definition 2.4.3(item 1))

iff $\forall F = (W, \mathscr{O}) \in \mathscr{F}, \forall M$ over F, $[\![\Gamma]\!] \subseteq [\![A]\!]$.

2.

$\Gamma \vDash^{g}_{\mathscr{F}} A$

iff $\forall F = (W, \mathscr{O}) \in \mathscr{F}, \forall M$ over F, [if $M \vDash \Gamma$ then $M \vDash A$]

(by definition of $\Gamma \vDash^{g}_{\mathscr{F}} A$, definition 2.6.1(item 2))

iff $\forall F = (W, \mathscr{O}) \in \mathscr{F}, \forall M$ over F, [if $M \in (\!|\Gamma|\!)_F$ then $M \in (\!|A|\!)_F$]

(by definition of $F \in (\!|\Gamma|\!)_F$, definition 2.6.5)

iff $\forall F = (W, \mathscr{O}) \in \mathscr{F}$, $(\!|\Gamma|\!)_F \subseteq (\!|A|\!)_F$.

□

We'll now spend some time discussing some useful theorems related to entailment.

The deduction theorem exposes the relationship between local entailment and our standard conditional connective. It is a very powerful theorem that will make our life much better at times.

Theorem 2.6.7 (Deduction Theorem (Semantic version))**.**
Given sentences A, B, a set of sentences, Γ, and a class of sphere frames $\mathscr{F}$,

$$\Gamma \cup \{ A \} \vDash^{\ell}_{\mathscr{F}} B \quad \textit{iff} \quad \Gamma \vDash^{\ell}_{\mathscr{F}} A \to B.$$

Proof. Let A and B be arbitrary sentences. Let Γ be an arbitrary set of sentences. Let $\mathscr{F}$ be a class of sphere frames.

For the first direction, suppose $\Gamma \cup \{ A \} \vDash^{\ell}_{\mathscr{F}} B$.

Let $F = (W, \mathscr{O}) \in \mathscr{F}$ be arbitrary, let M be an arbitrary model over F. Theorem 2.6.6 (item 1) tells us that $[\![\Gamma \cup \{ A \}]\!] \subseteq [\![B]\!]$. Note that $[\![\Gamma \cup \{ A \}]\!] = [\![\Gamma]\!] \cap [\![A]\!]$. We want to show that $[\![\Gamma]\!] \subseteq [\![A \to B]\!]$. Let $w \in [\![\Gamma]\!]$ be arbitrary. We have two cases:

Case 1: If $w \in [\![A]\!]$, then $w \in [\![\Gamma \cup \{ A \}]\!]$ and therefore $w \in [\![B]\!]$. Then $w \in [\![\neg A]\!] \cup [\![B]\!]$ and by definition $w \in [\![A \to B]\!]$.

Case 2: If $w \notin [\![A]\!]$, then $w \in [\![\neg A]\!] \cup [\![B]\!]$ and by definition $w \in [\![A \to B]\!]$.

Therefore, we've shown that $[\![\Gamma]\!] \subseteq [\![A \to B]\!]$ and since M and $F \in \mathscr{F}$ were both arbitrary, then we can apply theorem 2.6.6 (item 1) again we get that $\Gamma \vDash^{\ell}_{\mathscr{F}} A \to B$.

For the second direction, suppose $\Gamma \vDash^{\ell}_{\mathscr{F}} A \to B$. Let $F = (W, \mathscr{O}) \in \mathscr{F}$ be arbitrary, let M be an arbitrary model over F. Theorem 2.6.6 (item 1) tells us that $[\![\Gamma]\!] \subseteq [\![A \to B]\!]$. We want to show that $[\![\Gamma \cup \{ A \}]\!] \subseteq [\![B]\!]$. Let $w \in [\![\Gamma \cup \{ A \}]\!]$ be arbitrary. Then, $w \in [\![\Gamma]\!]$ and $w \in [\![A]\!]$. Since $w \in [\![\Gamma]\!]$ we have that $w \in [\![A \to B]\!]$. Then by definition, $w \in [\![\neg A]\!] \cup [\![B]\!]$, but since $w \in [\![A]\!]$ we must have that $w \in [\![B]\!]$. Therefore $[\![\Gamma \cup \{ A \}]\!] \subseteq [\![B]\!]$. Then since M and $F \in \mathscr{F}$ were arbitrary, we can apply theorem 2.6.6 (item 1) once again to get $\Gamma \cup \{ A \} \vDash^{\ell}_{\mathscr{F}} B$. □

We also have a simpler special case when Γ is the empty set.

Remark 2.6.8. When $\Gamma = \varnothing$ we get

$$A \vDash^{\ell}_{\mathscr{F}} B \quad \text{iff} \quad \vDash_{\mathscr{F}} A \to B.$$

The reader may wonder if there is a "global deduction theorem." Unfortunately there is not, it's one of the fundamental differences between global and local entailment.

Remark 2.6.9. There is no global version of the deduction theorem. To see why, consider our proof of theorem 2.6.4. We constructed an example where $\mathbb{P}_1 \vDash^g_{\mathscr{F}} \Box\mathbb{P}_1$ is true but $\mathbb{P}_1 \nvDash^\ell_{\mathscr{F}} \Box\mathbb{P}_1$. However, if we had a global deduction theorem, then could apply it to $\mathbb{P}_1 \vDash^g_{\mathscr{F}} \Box\mathbb{P}_1$ to obtain $\vDash_{\mathscr{F}} \mathbb{P}_1 \to \Box\mathbb{P}_1$. Then, using our local actual deduction theorem 2.6.7 we could obtain $\mathbb{P}_1 \vDash^\ell_{\mathscr{F}} \Box\mathbb{P}_1$ which contradicts with our counterexample.

Essentially, the issue comes down to the fact that global entailment deals with validity over entire models while local entailment deals with truth at individual worlds. It's almost like the difference between saying "if $\mathbb{P}_1$ is true everywhere, then..." as opposed to saying "if $\mathbb{P}_1$ is true here, then...."

The following two theorems will often be useful.

Theorem 2.6.10 (Monotony)**.**
Given a class of frames, $\mathscr{F}$, a set of sentences Γ, and a sentence A,

$$\textit{if } \Gamma' \subseteq \Gamma \textit{ and } \Gamma' \vDash^\ell_{\mathscr{F}} A \textit{ then } \Gamma \vDash^\ell_{\mathscr{F}} A.$$

Proof. Let $\mathscr{F}$ be a class of frames, Γ be a set of sentences, and A be a sentence. Suppose there is a subset $\Gamma' \subseteq \Gamma$ such that $\Gamma' \vDash^\ell_{\mathscr{F}} A$.

Let $F = (W, \mathscr{O}) \in \mathscr{F}$ be arbitrary, let M be an arbitrary model over F.

Then, by theorem 2.6.6(item 1), we have $[\![\Gamma']\!] \subseteq [\![A]\!]$. Since $\Gamma' \subseteq \Gamma$ then $[\![\Gamma]\!] = [\![\Gamma' \cup \Gamma]\!] = [\![\Gamma']\!] \cap [\![\Gamma]\!]$, meaning that $[\![\Gamma]\!] \subseteq [\![\Gamma']\!]$ and therefore $[\![\Gamma]\!] \subseteq [\![A]\!]$. Since M and $F \in \mathscr{F}$ are arbitrary, then by theorem 2.6.6 (item 1), we have $\Gamma \vDash^\ell_{\mathscr{F}} A$. □

Theorem 2.6.11 (Reflexivity)**.**
Given a class of frames, $\mathscr{F}$, a set of sentences Γ, and a sentence A,

$$\textit{If } A \in \Gamma \textit{ then } \Gamma \vDash^\ell_{\mathscr{F}} A.$$

Proof. Let A be a sentence, Γ be a set of sentences, and $\mathscr{F}$ be a class of sphere frames.

Suppose $A \in \Gamma$. Suppose M is an arbitrary model over an arbitrary $F \in \mathscr{F}$. We want to show that $[\![\Gamma]\!] \subseteq [\![A]\!]$. Note that since $A \in \Gamma$, then $\Gamma = \Gamma \cup \{ A \}$ and $[\![\Gamma]\!] = [\![\Gamma]\!] \cap [\![A]\!]$. Let $w \in [\![\Gamma]\!]$, then $w \in [\![A]\!]$. Since M and F were arbitrary, then by theorem 2.6.6 (item 1) we have $\Gamma \vDash^{\ell}_{\mathscr{F}} A$. □

2.7 Satisfiability

Earlier we described validity over a class of frames. A set of sentences are valid if they are true "everywhere" over a class of frames (i.e., at all worlds in all models over all frames in the class it is true). By negating the definition we obtain a notion of "invalidity" whereby a set of sentences is invalid if it is false "somewhere" (i.e., there exists a world in a model over a frame in the class where it's false). Taking these concepts further we can develop a notion of "satisfiability," where we say that a set of sentences is satisfiable if it's true "somewhere" and unsatisfiable if it's false "everywhere."

There's a close relationship between each of these semantic notions and in the following chapter we will see that they each have a syntactic counterpart. Later, as we develop the relationship between semantics and syntax it will at times be useful to switch between related notions.

For now we simply provide a definition and a few theorems that will be useful in later chapters.

Definition 2.7.1 (Satisfiable Set of Sentences)**.**
Given a class of frames $\mathscr{F}$, a set of sentences, Γ, is $\mathscr{F}$-satisfiable iff there exists a frame, $F \in \mathscr{F}$, and a model over $\mathscr{F}$, $M = (W, \mathscr{O}, v)$, and a world, $w \in W$, such that $M, w \vDash \Gamma$.

The following theorems expose some of the relationships between the notions of $\mathscr{F}$-satisfiability and entailment.

Theorem 2.7.2.
Given a class of frames, $\mathscr{F}$, and a set of sentences, Γ

$$\Gamma \text{ is } \mathscr{F}\text{-satisfiable} \quad \text{iff} \quad \Gamma \nvDash^{\ell}_{\mathscr{F}} \bot.$$

Proof. Let $\mathscr{F}$ be a class of frames and Γ be a set of sentences.

$$\begin{aligned}
&\Gamma \text{ is } \mathscr{F}\text{-satisfiable} \\
&\quad \text{iff } \exists F = (W, \mathscr{O}) \in \mathscr{F}, \exists M \text{ over } F, \exists w \in W, M, w \models \Gamma \\
&\quad \text{iff } \exists F = (W, \mathscr{O}) \in \mathscr{F}, \exists M \text{ over } F, \exists w \in W, \\
&\qquad (M, w \models \Gamma \text{ and } M, w \not\models \bot) \\
&\quad \text{iff } \Gamma \not\models^{\ell}_{\mathscr{F}} \bot.
\end{aligned}$$

□

Theorem 2.7.3.
Given a class of frames, $\mathscr{F}$, a set of sentences, Γ, and a sentence A,

$$\Gamma \not\models^{\ell}_{\mathscr{F}} A \quad \textit{iff} \quad \Gamma \cup \{ \neg A \} \textit{ is } \mathscr{F}\textit{-satisfiable}.$$

Proof. Let $\mathscr{F}$ be a class of frames, Γ be a set of sentences, and A be a sentence.

$$\begin{aligned}
&\Gamma \not\models^{\ell}_{\mathscr{F}} A \\
&\quad \text{iff } \exists F = (W, \mathscr{O}) \in \mathscr{F}, \exists M \text{ over } F, \exists w \in W, \\
&\qquad (M, w \models \Gamma \text{ and } M, w \not\models A) \\
&\quad \text{iff } \exists F = (W, \mathscr{O}) \in \mathscr{F}, \exists M \text{ over } F, \exists w \in W, \\
&\qquad (M, w \models \Gamma \text{ and } M, w \models \neg A) \\
&\quad \text{iff } \exists F = (W, \mathscr{O}) \in \mathscr{F}, \exists M \text{ over } F, \exists w \in W, \\
&\qquad M, w \models (\Gamma \cup \{ \neg A \}) \\
&\quad \text{iff } \Gamma \cup \{ \neg A \} \text{ is } \mathscr{F}\text{-satisfiable}.
\end{aligned}$$

□

Chapter 3

Syntax

In the previous chapter we described the *semantics* for our language by formulating models that assign truth values to our sentences and used them to reason about what it means for a sentences to be true or valid. While semantics are very useful, there is another equally useful way to reason about the sentences in our language. For this alternate approach we'll describe the *syntax* for our language which will precisely formalize the notions of derivations (i.e., proofs) and theorems within our logic. To do this, we'll take our language, equip it with a list of axioms and inference rules, and study the resulting sets of derivable sentences (which Lewis calls "V-logics").[1]

We'll start by introducing our inference rules and then explaining what it means to derive a sentence in our logic. Then we'll define a V-logic generally, as a set of sentences containing the axioms and closed under the set of inference rules. Our goal however, isn't to study V-logics in general, but rather V-logics generated by a set of axioms (which we'll call axiomatic systems). So, following Lewis, we'll then describe the smallest V-logic containing a set of axioms and we'll describe an axiomatic system generated by a set of axioms and show that the two notions are equivalent. This may seem roundabout, but by doing things this way we end up effectively saying that a V-logic is

[1]According to Lewis, the "V" stands for "variable strictness," in reference to how in the semantics, the sets of worlds one considers when evaluating a counterfactual statement may vary depending on the context.

exactly its set of theorems, and in doing so we intuitively address the issue that different axiomatizations can result in the same logic.

With the definition of a logic and derivation defined we'll develop a number of additional logical rules and prove a number of useful theorems. Next, we'll list a number of axiomatic systems of interest to Lewis and ourselves and prove a few things about them. Afterwards we'll extend our notion of derivability to derivability from a set of sentences and we'll have both local and global versions of this. With this in hand we can define a notion of consistency and prove Lindenbaum's lemma which will be useful in later chapters.

If it feels like some of these notions about syntax seem to suggestively parallel those about semantics, then rest assured that that is because there is a correspondence between the two which we'll make explicit in the Soundness and Completeness chapters.

3.1 Definition of a V-logic

Rules of inference tell us how we may obtain a sentence from axioms or previously obtained sentences. V-logics rely on two rules of inference which we'll define here.[2]

Definition 3.1.1 (Modus Ponens)**.**
Given sentences A and B, Modus Ponens tells us that we may obtain the sentence B from the sentences A and $A \to B$.

$$\frac{A \qquad A \to B}{B}\ \text{MP.}$$

Formally, we say B follows from A and $A \to B$ by MP.

Definition 3.1.2 (Comparative Possibility)**.**
Given a sentence A, and for $n \geq 1$, a set of sentences $\{ B_1, \dots, B_n \}$, Comparative Possibility tells us that we may obtain the sentence $(B_1 \preccurlyeq A) \vee \cdots \vee (B_n \preccurlyeq A)$ from the sentence $A \to (B_1 \vee \cdots \vee B_n)$.

[2]Since Lewis was particularly interested in the logic **VC** then he actually provides additional formulations of axioms and inference rules specifically for the V-logics **VC**, **VCS**, and **VW** in his book. Other formulations of these specific logics may be found in papers. We will not discuss those in this text since we're interested in more general logics.

$$\frac{A \to (B_1 \lor \cdots \lor B_n)}{(B_1 \preccurlyeq A) \lor \cdots \lor (B_n \preccurlyeq A)}\ \textsc{Cp}.$$

Formally, we say that $(B_1 \preccurlyeq A) \lor \cdots \lor (B_n \preccurlyeq A)$ follows from $A \to (B_1 \lor \cdots \lor B_n)$ by Cp.

Remark 3.1.3. In the case where $n = 1$, comparative possibility tells us that $B \preccurlyeq A$ follows from $A \to B$ by Cp.

Definition 3.1.4 (Derivation)**.**
A sentence A is **derivable from a set of sentences, Σ,** if and only if there exists a finite sequence of sentences $A_1, A_2, \ldots, A_n$ where $A_n = A$ and for each $i \in \{ 1, \ldots, n \}$, either

1. A_i is an instance of an axiom, where our axioms are:
   a) All classical propositional tautologies definition 2.1.11.
   b) **Transitivity:** $[(p \preccurlyeq q) \land (q \preccurlyeq r)] \to (p \preccurlyeq r)$, (where $p, q, r \in$ Props, see remark 2.1.3).
   c) **Strong Connectivity:** $(p \preccurlyeq q) \lor (q \preccurlyeq p)$, (where $p, q \in$ Props, see remark 2.1.3).
   d) All sentence in Σ.
2. A_i follows from previous sentences by an inference rule. More precisely, either A_i follows from:
   a) sentences A_j and A_k, where $j, k < i$, by Mp, or
   b) a sentence A_j, where $j < i$, by Cp.

If A is derivable from a set of sentences, Σ, we say that the sequence of sentences, $A_1, \ldots, A_n$, is a derivation of A. Moreover, we write:

$$\vdash_\Sigma A,$$

if such a derivation exists.

When performing a derivation we write out the sequence of sentences explicitly as a numbered list. Each sentence is annotated differently depending on whether it is an instance of an axiom or it follows from a rule. Instances of axioms are annotated: Taut if it is an instance of a classical tautology, Trans or Connex if it is an instance of transitivity or strong connectivity respectively, or by name if it is an instance of a sentence in Σ. Any classical propositional tautologies used will be listed below the derivation. If the sentence follows from a previous sentence by either Mp or Cp then the rule will be written along with the corresponding line numbers for the sentences it follows from. For convenience we will also number sentences that are obtained by applying the definition of a defined connective, though technically speaking this does not introduce a new sentence into our sequence. Later we will derive more rules which we will also use and annotate similarly.

Example 3.1.5.
Let $\mathbf{W} := [(\Box\, p \vee \boxdot\, p) \to p]$. Suppose

$$\Sigma = \{\, \mathbf{W} \,\}\,.$$

We will show that $\vdash_\Sigma \Box\, A \to A$. Recall the lowercase p in the definition of $\mathbf{W}$ represents an atomic sentence, while the uppercase A here represents an arbitrary sentence (obtained by instantiation).

1. $\vdash_\Sigma \Box\, A \to (\Box\, A \vee \boxdot\, A)$ — Taut
2. $\vdash_\Sigma (\Box\, A \vee \boxdot\, A) \to A$ — $\mathbf{W}$
3. $\vdash_\Sigma [\Box\, A \to (\Box\, A \vee \boxdot\, A)] \to ([(\Box\, A \vee \boxdot\, A) \to A] \to [\Box\, A \to A])$ — Taut
4. $\vdash_\Sigma [(\Box\, A \vee \boxdot\, A) \to A] \to [\Box\, A \to A]$ — Mp, 1,3
5. $\vdash_\Sigma \Box\, A \to A$ — Mp, 2,4

Lines 1 and 3 are instances of the following classical propositional tautologies:

$$p \to (p \vee q)$$

$$(p \to q) \to [(q \to r) \to (p \to r)].$$

Moreover, since $(\Box A \to A)$ is an arbitrary instance of the sentence $\mathbf{T} := (\Box p \to p)$ then we've actually shown that every instance of the sentence $\mathbf{T}$ is derivable from $\mathbf{W}$. We'll explore this notion more in-depth later.

Definition 3.1.6 (V-Logic)**.**
A set of sentences, Λ, is a **V-Logic** if and only if:

1. Λ contains all instances of:
   a) all classical propositional tautologies definition 2.1.11.
   b) **Transitivity:** $[(p \preccurlyeq q) \wedge (q \preccurlyeq r)] \to (p \preccurlyeq r)$, where (where $p, q, r \in$ Props, see remark 2.1.3).
   c) **Strong Connectivity:** $(p \preccurlyeq q) \vee (q \preccurlyeq p)$, (where $p, q \in$ Props, see remark 2.1.3).
2. Λ is closed under the inference rules:
   a) Modus Ponens: For all sentences, A and B,
   $$\text{if } A \in \Lambda \text{ and } (A \to B) \in \Lambda \text{ then } B \in \Lambda.$$
   b) Comparative Possibility: For all $n \geq 1 \in \mathbb{N}$, for all sentences A and $B_1, \ldots, B_n$,
   $$\text{if } (A \to (B_1 \vee \cdots \vee B_n)) \in \Lambda \text{ then } ((B_1 \preccurlyeq A) \vee \cdots \vee (B_n \preccurlyeq A)) \in \Lambda.$$

We'll sometimes simply refer to these as **logics**.

Theorem 3.1.7.
The set of all sentences, L_V, is a V-logic. In particular, it is the largest V-Logic.

Proof. This set trivially satisfies the definition of a V-logic. □

Theorem 3.1.8.
Given a set of sentences, $\Sigma = \{ A_1, A_2, \ldots, A_n \}$, there exists a smallest V-logic, Λ_Σ, such that Λ_Σ contains all instances of all sentences in Σ.

Proof. Let $\Sigma = \{ A_1, A_2, \ldots, A_n \}$ be a set of sentences.

Consider the set of V-logics

$$\{ \Lambda \mid \Lambda \text{ contains all instances of all sentences in } \Sigma \text{ and } \Lambda \text{ is a V-logic} \}.$$

Since L_V is a V-logic containing all sentences, then the set is non-empty. Take the intersection,

$$\Lambda_\Sigma := \bigcap \left\{ \Lambda \,\middle|\, \begin{array}{l} \Lambda \text{ contains all instances of all sentences in } \Sigma \text{ and} \\ \quad \Lambda \text{ is a V-logic} \end{array} \right\}.$$

To prove that Λ_Σ is indeed a V-logic we need to show that Λ_Σ contains all instances of classical propositional tautologies, transitivity, and strong connectivity (which it does so trivially) and that it's closed under the inference rules Mp and Cp. If it were not closed under a rule then there would be at least one V-logic in that set which would also not be closed under said rule but that would be a contradiction. Next, to prove that all instances of all sentences in Σ are contained in Λ_Σ, we just need to observe that all instances of all sentences in Σ are contained in every logic in that set. □

Theorem 3.1.9.
Given a set of sentences, $\Sigma = \{ A_1, A_2, \ldots, A_n \}$, the set of sentences derivable from Σ is equal to the smallest logic containing all instances of sentences in Σ, Λ_Σ, i.e.,

$$\vdash_\Sigma A \textit{ iff } A \in \Lambda_\Sigma.$$

Proof. Suppose $\vdash_\Sigma A$. Then there is a sequence of sentences $A_1, \ldots, A_n$, such that $A_n = A$ and for each A_i, it is either an instance of a classical propositional tautology, transitivity, strong connectivity, or of a sentence in Σ; or it follows from previous sentences in the sequence by either Mp or Cp. In every case, it's clear that each A_i must be in Λ_Σ and therefore $A \in \Lambda_\Sigma$.

Suppose $A \in \Lambda_\Sigma$. Recall that Λ_Σ is the smallest logic containing all instances of Σ. It follows, that either A is an instance of a classical propositional tautology, transitivity, strong connectivity, or of a sentence in Σ; or there are other sentences in Λ_Σ from which A follows by Mp or Cp. By iterating on this reasoning we can obtain a finite derivation of A showing that $\vdash_\Sigma A$. □

Definition 3.1.10 (Axiomatic System)**.**
Given a set of sentences,

$$\Sigma = \{ A_1, A_2, \dots, A_n \},$$

we call the smallest logic containing all instances of sentences in Σ, Λ_Σ, the **V-logic generated by** Σ or the **axiomatic system** Σ. Additionally, the sentences in Σ may be called **axioms** of the axiomatic system alongside instances of transitivity, connectivity, and classical propositional tautologies.

Moreover, a sentence A is called a **theorem** of Σ iff $\vdash_\Sigma A$.

Instead of writing Σ we may write $\mathbf{VA_1A_2 \cdots A_n}$. For instance, if $\Sigma = \{ \mathbf{T}, \mathbf{4} \}$ where **T** and **4** are sentences, then we may refer to the axiomatic system **VT4** and write $\vdash_{\mathbf{VT4}} A$ to refer to its theorems.

If $\Sigma = \varnothing$, then Λ_Σ is called the axiomatic system **V**. This axiomatic system contains all instances of classical tautologies, transitivity, and strong connectivity. It is the smallest V-logic and every theorem of **V** is a theorem of every V-logic.

3.2 Additional Rules

In order to make it more convenient to perform derivations we will derive a few rules that are not included in Lewis (2001).

Due to the way derivations are defined, if we want to perform a simple manipulation that follows from the rules of classical propositional logic we often have to resort to complicated instantiations and applications of modus ponens. This rule will allow us to perform any manipulation that follows from previous sentences via classical propositional logic effortlessly.

Theorem 3.2.1 (The Rule Pl)**.**
Given an axiomatic system, Σ*, a sentence* B*, sentences* $A_1, A_2, \dots, A_n$*.*

If $\vdash_\Sigma A_1, \quad \vdash_\Sigma A_2, \quad \cdots, \quad \vdash_\Sigma A_n$
and B *follows from* $A_1, A_2, \dots, A_n$ *by classical propositional logic*
then $\vdash_\Sigma B$.

Proof. Let Σ be an axiomatic system and let $A_1, A_2, \ldots, A_n, B$ be sentences. Suppose we have derived,

$$\vdash_\Sigma A_1, \quad \vdash_\Sigma A_2, \quad \cdots, \quad \vdash_\Sigma A_n,$$

and that B follows from $A_1, A_2, \ldots, A_n$ by classical propositional logic.

Since B follows from $A_1, A_2, \ldots, A_n$ by classical propositional logic, then we have the theorem $\vdash_\Sigma A_1 \to (A_2 \to \cdots (A_n \to B) \cdots)$ and by applying MP n times with each corresponding derived sentence we may obtain $\vdash_\Sigma B$.

We'll annotate this rule with PL and a line number for each derived sentence from which B follows by classical propositional logic. □

At times we may have a complicated sentence containing A, and assuming we've derived $\vdash_\Sigma A \leftrightarrow B$, we may want to substitute A with B. Currently this is a very tedious thing to do, so we'll introduce a rule to remedy that. Recall our notation for substitution remark 2.1.10.

Theorem 3.2.2 (Substitution Theorem)**.**
Given an axiomatic system, Σ, sentences A, B, and C, and atomic sentence $p \in$ PROPS,

$$\text{if } \vdash_\Sigma A \leftrightarrow B \text{ then } \vdash_\Sigma C[p/A] \leftrightarrow C[p/B].$$

Proof. Let Σ be an axiomatic system and let A, B, C be sentences.

Suppose $\vdash_\Sigma A \leftrightarrow B$. We wish to prove that $\vdash_\Sigma C[p/A] \leftrightarrow C[p/B]$.

The proof is by induction on the complexity of C.

Base cases:

1. If $C := \bot$

1.	$\vdash_\Sigma \bot \leftrightarrow \bot$	TAUT
2.	$\vdash_\Sigma C[p/A] \leftrightarrow C[p/B]$	Defn of $C[p/A]$ and $C[p/B]$

2. If $C := q$ (where $q \neq p$)

1.	$\vdash_\Sigma q \leftrightarrow q$	TAUT
2.	$\vdash_\Sigma C[p/A] \leftrightarrow C[p/B]$	Defn of $C[p/A]$ and $C[p/B]$

3. If $C := p$

1.	$\vdash_\Sigma A \leftrightarrow B$	By Assumption
2.	$\vdash_\Sigma C[p/A] \leftrightarrow C[p/B]$	Defn of $C[p/A]$ and $C[p/B]$

Inductive Hypothesis: Let F and G be sentences, such that

$$\vdash_\Sigma F[p/A] \leftrightarrow F[p/B] \text{ and } \vdash_\Sigma G[p/A] \leftrightarrow G[p/B].$$

Inductive Step:

1. If $C := F \to G$, then $C[p/A] = F[p/A] \to G[p/A]$ and $C[p/B] = F[p/B] \to G[p/B]$. We wish to show $\vdash_\Sigma C[p/A] \leftrightarrow C[p/B]$. Then

1.	$\vdash_\Sigma F[p/A] \leftrightarrow F[p/B]$	By IH
2.	$\vdash_\Sigma G[p/A] \leftrightarrow G[p/B]$	By IH
3.	$\vdash_\Sigma (F[p/A] \leftrightarrow F[p/B])$	
	$\to [(F[p/A] \to G[p/A]) \leftrightarrow (F[p/B] \to G[p/A])]$	Taut
4.	$\vdash_\Sigma (F[p/A] \to G[p/A]) \leftrightarrow (F[p/B] \to G[p/A])$	Mp,1,3
5.	$\vdash_\Sigma (G[p/A] \leftrightarrow G[p/B])$	
	$\to [(F[p/B] \to G[p/A]) \leftrightarrow (F[p/B] \to G[p/B])]$	Taut
6.	$\vdash_\Sigma (F[p/B] \to G[p/A]) \leftrightarrow (F[p/B] \to G[p/B])$	Mp,2,5
7.	$\vdash_\Sigma (F[p/A] \to G[p/A]) \leftrightarrow (F[p/B] \to G[p/B])$	Pl,4,6
8.	$\vdash_\Sigma C[p/A] \leftrightarrow (F[p/B] \to G[p/B])$	Defn $C[p/A]$
9.	$\vdash_\Sigma C[p/A] \leftrightarrow C[p/B]$	Defn $C[p/B]$

Classical tautologies used:

$$(p \leftrightarrow q) \to [(q \to r) \leftrightarrow (p \to r)]$$
$$(p \leftrightarrow q) \to [(r \to p) \leftrightarrow (r \to q)].$$

2. If $C := F \preccurlyeq G$, then $C[p/A] = F[p/A] \preccurlyeq G[p/A]$ and $C[p/B] = F[p/B] \preccurlyeq G[p/B]$. We wish to show $\vdash_\Sigma C[p/A] \leftrightarrow C[p/B]$.

   Then

1.	$\vdash_\Sigma F[p/A] \leftrightarrow F[p/B]$	By IH
2.	$\vdash_\Sigma F[p/B] \approx F[p/A]$	Cp,1, twice
3.	$\vdash_\Sigma G[p/A] \leftrightarrow G[p/B]$	By IH
4.	$\vdash_\Sigma G[p/B] \approx G[p/A]$	Cp,3, twice
5.	$\vdash_\Sigma [(F[p/B] \preccurlyeq F[p/A]) \wedge (F[p/A] \preccurlyeq G[p/A])]$ $\rightarrow (F[p/B] \preccurlyeq G[p/A])$	Trans
6.	$\vdash_\Sigma (F[p/B] \preccurlyeq F[p/A])$ $\rightarrow [(F[p/A] \preccurlyeq G[p/A]) \rightarrow (F[p/B] \preccurlyeq G[p/A])]$	Pl,5
7.	$\vdash_\Sigma [(F[p/A] \preccurlyeq F[p/B]) \wedge (F[p/B] \preccurlyeq G[p/A])]$ $\rightarrow (F[p/A] \preccurlyeq G[p/A])$	Trans
8.	$\vdash_\Sigma (F[p/A] \preccurlyeq F[p/B])$ $\rightarrow [(F[p/B] \preccurlyeq G[p/A]) \rightarrow (F[p/A] \preccurlyeq G[p/A])]$	Pl,7
9.	$\vdash_\Sigma (F[p/B] \approx F[p/A])$ $\rightarrow [(F[p/A] \preccurlyeq G[p/A]) \leftrightarrow (F[p/B] \preccurlyeq G[p/A])]$	Pl,6,8
10.	$\vdash_\Sigma (F[p/A] \preccurlyeq G[p/A]) \leftrightarrow (F[p/B] \preccurlyeq G[p/A])$	Mp,2,8
11.	$\vdash_\Sigma [(F[p/B] \preccurlyeq G[p/A]) \wedge (G[p/A] \preccurlyeq G[p/B])]$ $\rightarrow (F[p/B] \preccurlyeq G[p/B])$	Trans
12.	$\vdash_\Sigma [(F[p/B] \preccurlyeq G[p/B]) \wedge (G[p/B] \preccurlyeq G[p/A])]$ $\rightarrow (F[p/B] \preccurlyeq G[p/A])$	Trans
13.	$\vdash_\Sigma (F[p/B] \preccurlyeq G[p/A]) \leftrightarrow (F[p/B] \preccurlyeq G[p/B])$	Pl,4,11,12
14.	$\vdash_\Sigma (F[p/A] \preccurlyeq G[p/A]) \leftrightarrow (F[p/B] \preccurlyeq G[p/B])$	Pl,10,13
15.	$\vdash_\Sigma C[p/A] \leftrightarrow (F[p/B] \preccurlyeq G[p/B])$	Defn $C[p/A]$
16.	$\vdash_\Sigma C[p/A] \leftrightarrow C[p/B]$	Defn $C[p/B]$

This concludes our induction. □

We now obtain the following rule that lets us swap occurrences of the sentence A with the sentence B, provided we have derived $\vdash_\Sigma A \leftrightarrow B$ and $\vdash_\Sigma C[p/A]$.

Theorem 3.2.3 (The rule Eq)**.**

$$\textit{If } \vdash_\Sigma A \leftrightarrow B \textit{ and } \vdash_\Sigma C[p/A] \textit{ then } \vdash_\Sigma C[p/B].$$

Proof. Result follows by combining the substitution theorem with Mp. □

We'll annotate this rule with Eq and two line numbers, one for the sentence we're manipulating and the other for the line number deriving $\vdash_\Sigma \boldsymbol{A} \leftrightarrow \boldsymbol{B}$.

Theorem 3.2.4.
Given an axiomatic system Σ, if $\vdash_\Sigma \boldsymbol{A}$ and if $\boldsymbol{B}$ is an instance of $\boldsymbol{A}$, then $\vdash_\Sigma \boldsymbol{B}$.

Proof. If $\vdash_\Sigma \boldsymbol{A}$ then we have a derivation $A_1, \dots, A_n$ of $\boldsymbol{A}$. We then simply need to apply the same substitution that we applied to go from $\boldsymbol{A}$ to $\boldsymbol{B}$ to every sentence in the derivation. Then simply observe that an instance of an instance is an instance and the inference rules are closed under instantiation. Therefore $\vdash_\Sigma \boldsymbol{B}$.

A more precise but tedious proof may be performed by induction on the length of derivation of $\boldsymbol{A}$. □

Remark 3.2.5. In a traditional proof by contradiction where we want to show that $\boldsymbol{B}$ follows from $(A_1 \wedge A_2 \wedge \cdots \wedge A_n)$, we proceed by assuming $A_1, A_2, \dots, A_n$, then supposing $\neg \boldsymbol{B}$ (for the sake of contradiction), and finally obtaining a contradiction. Effectively, the idea is that we prove a statement of the form $(A_1 \wedge A_2 \wedge \cdots \wedge A_n) \rightarrow (\neg \boldsymbol{B} \rightarrow \bot)$.

In an axiomatic system, Σ, we may achieve something similar to proof by contradiction by using the rule Eq. The strategy is to start with an instance of the classical tautology, $p \rightarrow (q \rightarrow (p \wedge q))$, and produce a derivation of the following form:

1.	$\vdash_\Sigma (A_1 \wedge A_2 \wedge \cdots \wedge A_n) \rightarrow$	
	$(\neg \boldsymbol{B} \rightarrow (A_1 \wedge A_2 \wedge \cdots \wedge A_n \wedge \neg \boldsymbol{B}))$	Taut
⋮	⋮	⋮
k.	$\vdash_\Sigma (A_1 \wedge A_2 \wedge \cdots \wedge A_n \wedge \neg \boldsymbol{B}) \leftrightarrow \bot$	(?)
k+1.	$\vdash_\Sigma (A_1 \wedge A_2 \wedge \cdots \wedge A_n) \rightarrow (\neg \boldsymbol{B} \rightarrow \bot)$	Eq,1,2
k+2.	$\vdash_\Sigma (A_1 \wedge A_2 \wedge \cdots \wedge A_n) \rightarrow \boldsymbol{B}$	Pl,3

The goal is to somehow derive the line numbered, (k). With that done the rest of the argument falls into place.

We can also derive the *Necessitation* rule from normal modal logic for our outer modality.

Theorem 3.2.6 (The Rule Nec)**.**
Given an axiomatic system Σ*, for any sentence* A*,*

$$\textit{if } \vdash_\Sigma A, \textit{ then } \vdash_\Sigma \Box A.$$

Proof.

1.	$\vdash_\Sigma A$	Assumption
2.	$\vdash_\Sigma \neg A \to \bot$	Pl,1
3.	$\vdash_\Sigma \bot \preccurlyeq \neg A$	Cp,2
4.	$\vdash_\Sigma \Box A$	Defn of $\Box$

□

This rule is annotated with Nec and a line number of the sentence it's being applied to.

In addition, we obtain a similar rule for the inner modalities. We call this one *Possibilitation*.

Theorem 3.2.7 (The Rule Pos)**.**
Given an axiomatic system Σ*, for any sentence* A*,*

$$\textit{if } \vdash_\Sigma A, \textit{ then } \vdash_\Sigma \Diamonddot A.$$

Proof.

1.	$\vdash_\Sigma A$	Assumption
2.	$\vdash_\Sigma \top \to A$	Pl,1
3.	$\vdash_\Sigma A \preccurlyeq \top$	Cp,2
4.	$\vdash_\Sigma \Diamonddot A$	Defn of $\Diamonddot$

□

This rule is annotated with Pos and a line number of the sentence it's being applied to.

3.3 Derivations

In *Counterfactuals* (Lewis, 2001), Lewis defines $\Box A$ as $\bot \approx \neg A$ and $\Diamond A$ as $A \approx \top$. Since we define them differently in definition 2.1.6, we will take the time to show that both pairs of definitions are equivalent in our syntax.

Theorem 3.3.1.
Given an axiomatic system, Σ. The following are true for all sentences A.

1. $\vdash_\Sigma \neg A \preccurlyeq \bot$.

2. $\vdash_\Sigma \top \preccurlyeq A$.

3. $\vdash_\Sigma (\bot \approx \neg A) \leftrightarrow \Box A$.

4. $\vdash_\Sigma (A \approx \top) \leftrightarrow \Diamond A$.

Proof. Let Σ be an axiomatic system and let A be an arbitrary sentence.

1. We provide a derivation of $\vdash_\Sigma \neg A \preccurlyeq \bot$.

1.	$\vdash_\Sigma \bot \to \neg A$	Taut
2.	$\vdash_\Sigma \neg A \preccurlyeq \bot$	Cp,1

2. We provide a derivation of $\vdash_\Sigma \top \preccurlyeq A$.

1.	$\vdash_\Sigma A \to \top$	Taut
2.	$\vdash_\Sigma \top \preccurlyeq A$	Cp,1

3. We provide a derivation of $\vdash_\Sigma (\bot \approx \neg A) \leftrightarrow \Box A$.

1.	$\vdash_\Sigma (\neg A \preccurlyeq \bot) \to$ $([\bot \preccurlyeq \neg A] \to$ $[(\bot \preccurlyeq \neg A) \wedge (\neg A \preccurlyeq \bot)])$	Taut
2.	$\vdash_\Sigma \neg A \preccurlyeq \bot$	By 3.3.1(1)
3.	$\vdash_\Sigma [\bot \preccurlyeq \neg A] \to [(\bot \preccurlyeq \neg A) \wedge (\neg A \preccurlyeq \bot)]$	Mp,1,2
4.	$\vdash_\Sigma [(\bot \preccurlyeq \neg A) \wedge (\neg A \preccurlyeq \bot)] \to [\bot \preccurlyeq \neg A]$	Taut
5.	$\vdash_\Sigma [(\bot \preccurlyeq \neg A) \wedge (\neg A \preccurlyeq \bot)] \leftrightarrow [\bot \preccurlyeq \neg A]$	Pl,3,4
6.	$\vdash_\Sigma (\bot \approx \neg A) \leftrightarrow [\bot \preccurlyeq \neg A]$	Defn of $\approx$
7.	$\vdash_\Sigma (\bot \approx \neg A) \leftrightarrow \Box A$	Defn of $\Box$

Classical tautologies used

$$p \to (q \to [p \wedge q])$$

$$(p \wedge q) \to p.$$

4. We provide a derivation of $\vdash_\Sigma (A \approx \top) \leftrightarrow \diamond A$.

1.	$\vdash_\Sigma (\top \preccurlyeq A) \to$ $([A \preccurlyeq \top] \to$ $[(A \preccurlyeq \top) \wedge (\top \preccurlyeq A)])$	Taut
2.	$\vdash_\Sigma \top \preccurlyeq A$	By 3.3.1(2)
3.	$\vdash_\Sigma [A \preccurlyeq \top] \to [(A \preccurlyeq \top) \wedge (\top \preccurlyeq A)]$	Mp,1,2
4.	$\vdash_\Sigma [(A \preccurlyeq \top) \wedge (\top \preccurlyeq A)] \to [A \preccurlyeq \top]$	Taut
5.	$\vdash_\Sigma [(A \preccurlyeq \top) \wedge (\top \preccurlyeq A)] \leftrightarrow [A \preccurlyeq \top]$	Pl,3,4
6.	$\vdash_\Sigma (A \approx \top) \leftrightarrow [A \preccurlyeq \top]$	Defn of $\approx$
7.	$\vdash_\Sigma (A \approx \top) \leftrightarrow \diamond A$	Defn of $\diamond$

Classical tautologies used

$$p \to (q \to [p \wedge q])$$

$$(p \wedge q) \to p.$$

$\square$

The following theorems tell us that $\prec$ is a strict partial order and give us some other alternate versions of transitivity.

Theorem 3.3.2.
Given an axiomatic system, Σ, we have the following theorems in Σ:

1. $\preccurlyeq$ *is reflexive,*

   $\vdash_\Sigma A \preccurlyeq A.$

2. $\prec$ *is irreflexive,*

   $\vdash_\Sigma \neg(A \prec A).$

3. $\prec$ *can be strengthened to* $\preccurlyeq$*,*

   $\vdash_\Sigma (A \prec B) \rightarrow (A \preccurlyeq B).$

4. $\prec$ *is asymmetric,*

   $\vdash_\Sigma (A \prec B) \rightarrow \neg(B \prec A).$

5. $\vdash_\Sigma ((A \prec B) \wedge (B \preccurlyeq C)) \rightarrow (A \prec C).$

6. $\vdash_\Sigma ((A \preccurlyeq B) \wedge (B \prec C)) \rightarrow (A \prec C).$

7. $\prec$ *is transitive,*

   $\vdash_\Sigma ((A \prec B) \wedge (B \prec C)) \rightarrow (A \prec C).$

Proof. Let Σ be a V-Logic. We prove each statement separately:

1. We provide a derivation of $\vdash_\Sigma A \preccurlyeq A$.

1.	$\vdash_\Sigma (A \preccurlyeq A) \vee (A \preccurlyeq A)$	Connex
2.	$\vdash_\Sigma A \preccurlyeq A$	Pl,1

2. We provide a derivation of $\vdash_\Sigma \neg(A \prec A)$.

1.	$\vdash_\Sigma A \preccurlyeq A$	By 3.3.2(1)
2.	$\vdash_\Sigma \neg(A \prec A)$	Defn of $\prec$

3. We provide a derivation of $\vdash_\Sigma (A \prec B) \to (A \preccurlyeq B)$.

1.	$\vdash_\Sigma (B \preccurlyeq A) \vee (A \preccurlyeq B)$	Connex
2.	$\vdash_\Sigma \neg(B \preccurlyeq A) \to (A \preccurlyeq B)$	Pl,1
3.	$\vdash_\Sigma (A \prec B) \to (A \preccurlyeq B)$	Defn of $\prec$

4. We provide a derivation of $\vdash_\Sigma (A \prec B) \to \neg(B \prec A)$.

1.	$\vdash_\Sigma (A \prec B) \to (A \preccurlyeq B)$	By 3.3.2(3)
2.	$\vdash_\Sigma (A \prec B) \to \neg(B \prec A)$	Defn of $\prec$

5. We provide a derivation of $\vdash_\Sigma ((A \prec B) \wedge (B \preccurlyeq C)) \to (A \prec C)$.

1.	$\vdash_\Sigma [(A \prec B) \wedge (B \preccurlyeq C)] \to$ $[(C \preccurlyeq A) \to [(A \prec B) \wedge (B \preccurlyeq C) \wedge (C \preccurlyeq A)]]$	Taut
2.	$\vdash_\Sigma [(B \preccurlyeq C) \wedge (C \preccurlyeq A)] \to (B \preccurlyeq A)$	Trans
3.	$\vdash_\Sigma [(A \prec B) \wedge (B \preccurlyeq C) \wedge (C \preccurlyeq A)] \to$ $[(A \prec B) \wedge (B \preccurlyeq A)]$	Pl,2
4.	$\vdash_\Sigma [(A \prec B) \wedge (B \preccurlyeq C) \wedge (C \preccurlyeq A)] \to$ $[\neg(B \preccurlyeq A) \wedge (B \preccurlyeq A)]$	Defn of $\prec$
5.	$\vdash_\Sigma [(A \prec B) \wedge (B \preccurlyeq C) \wedge (C \preccurlyeq A)] \to \bot$	Pl,4
6.	$\vdash_\Sigma [(A \prec B) \wedge (B \preccurlyeq C)] \to [(C \preccurlyeq A) \to \bot]$	Pl,1,5
7.	$\vdash_\Sigma [(A \prec B) \wedge (B \preccurlyeq C)] \to (A \prec C)$	Pl,6

Classical tautology used:

$$p \to (q \to (p \wedge q)).$$

6. We provide a derivation of $\vdash_\Sigma ((A \preccurlyeq B) \wedge (B \prec C)) \to (A \prec C)$.

1.	$\vdash_\Sigma [(A \preccurlyeq B) \wedge (B \prec C)] \rightarrow$	
	$[(C \preccurlyeq A) \rightarrow [(A \preccurlyeq B) \wedge (B \prec C) \wedge (C \preccurlyeq A)]]$	Taut
2.	$\vdash_\Sigma [(C \preccurlyeq A) \wedge (A \preccurlyeq B)] \rightarrow (C \preccurlyeq B)$	Trans
3.	$\vdash_\Sigma [(A \preccurlyeq B) \wedge (B \prec C) \wedge (C \preccurlyeq A)] \rightarrow$	
	$[(B \prec C) \wedge (C \preccurlyeq B)]$	Pl,2
4.	$\vdash_\Sigma [(A \preccurlyeq B) \wedge (B \prec C) \wedge (C \preccurlyeq A)] \rightarrow$	
	$[\neg(C \preccurlyeq B) \wedge (C \preccurlyeq B)]$	Defn of $\prec$
5.	$\vdash_\Sigma [(A \preccurlyeq B) \wedge (B \prec C) \wedge (C \preccurlyeq A)] \rightarrow \bot$	Pl,4
6.	$\vdash_\Sigma [(A \preccurlyeq B) \wedge (B \prec C)] \rightarrow [(C \preccurlyeq A) \rightarrow \bot]$	Pl,1,5
7.	$\vdash_\Sigma [(A \preccurlyeq B) \wedge (B \prec C)] \rightarrow (A \prec C)$	Pl,6

Classical tautology used:

$$p \rightarrow (q \rightarrow (p \wedge q)).$$

7. We provide a derivation of $\vdash_\Sigma ((A \prec B) \wedge (B \prec C)) \rightarrow (A \prec C)$.

1.	$\vdash_\Sigma [(A \prec B) \wedge (B \preccurlyeq C)] \rightarrow (A \prec C)$	By 3.3.2(5)
2.	$\vdash_\Sigma (B \prec C) \rightarrow (B \preccurlyeq C)$	By 3.3.2(3)
3.	$\vdash_\Sigma [(A \prec B) \wedge (B \prec C)] \rightarrow [(A \prec B) \wedge (B \preccurlyeq C)]$	Pl,2
4.	$\vdash_\Sigma [(A \prec B) \wedge (B \prec C)] \rightarrow (A \prec C)$	Pl,1,3

□

We may also be interested in showing that the Kripke axiom, from modal logic, holds for our outer modality. To do that we will need this lemma.

Lemma 3.3.3.

$$\vdash_\Sigma (C \preccurlyeq (A \vee B)) \rightarrow ((C \preccurlyeq A) \wedge (C \preccurlyeq B))$$

Proof.

1.	$\vdash_\Sigma A \to (A \lor B)$	Taut
2.	$\vdash_\Sigma (A \lor B) \preccurlyeq A$	Cp,1
3.	$\vdash_\Sigma [(C \preccurlyeq (A \lor B)) \land ((A \lor B) \preccurlyeq A)] \to (C \preccurlyeq A)$	Trans
4.	$\vdash_\Sigma [(A \lor B) \preccurlyeq A] \to [(C \preccurlyeq (A \lor B)) \to (C \preccurlyeq A)]$	Pl,3
5.	$\vdash_\Sigma (C \preccurlyeq (A \lor B)) \to (C \preccurlyeq A)$	Mp, 2,4
6.	$\vdash_\Sigma B \to (A \lor B)$	Taut
7.	$\vdash_\Sigma (A \lor B) \preccurlyeq B$	Cp,1
8.	$\vdash_\Sigma [(C \preccurlyeq (A \lor B)) \land ((A \lor B) \preccurlyeq B)] \to (C \preccurlyeq B)$	Trans
9.	$\vdash_\Sigma [(A \lor B) \preccurlyeq B] \to [(C \preccurlyeq (A \lor B)) \to (C \preccurlyeq B)]$	Pl,8
10.	$\vdash_\Sigma (C \preccurlyeq (A \lor B)) \to (C \preccurlyeq B)$	Mp, 7,9
11.	$\vdash_\Sigma (C \preccurlyeq (A \lor B)) \to [(C \preccurlyeq A) \land (C \preccurlyeq B)]$	Pl, 5,10

□

Theorem 3.3.4 (Kripke Axiom)**.**

$$\vdash_\Sigma \Box(A \land B) \to (\Box A \land \Box B)$$

Proof.

1.	$\vdash_\Sigma (\neg A \lor \neg B) \leftrightarrow \neg(A \land B)$	Taut
2.	$\vdash_\Sigma [\bot \preccurlyeq (\neg A \lor \neg B)] \to ([\bot \preccurlyeq \neg A] \land [\bot \preccurlyeq \neg B])$	By 3.3.3
3.	$\vdash_\Sigma [\bot \preccurlyeq \neg(A \land B)] \to ([\bot \preccurlyeq \neg A] \land [\bot \preccurlyeq \neg B])$	Eq,1,2
4.	$\vdash_\Sigma \Box(A \land B) \to (\Box A \land \Box B)$	Defn of $\Box$

□

We'll see in the next section that we also have an inner version of the Kripke axiom but it only holds in certain axiomatic systems, not in general.

3.4 Axioms

Lewis (2001) introduces a variety of axioms which he then uses to describe a family of 26 axiomatic systems. In this section we'll present those axioms and the most important resulting axiomatic systems, and then we'll present some derivations within those axiomatic systems showing that some axioms can be used to derive others (and therefore are stronger).

Definition 3.4.1 (Axioms)**.**
The axioms presented by Lewis are reproduced here:[3]

1. $\mathbf{N} := \top \prec \bot$
2. $\mathbf{T} := \Box A \to A$
3. $\mathbf{W} := (\Box A \vee \boxdot A) \to A$
4. $\mathbf{C} := \diamonddot A \to A$
5. $\mathbf{S} := ((A \wedge B) \approx (A \wedge \neg B)) \to \neg \Diamond A$
6. $\mathbf{4} := \Box A \to \Box\Box A$
7. $\mathbf{5} := \Diamond A \to \Box \Diamond A$
8. $\mathbf{A}_1 := (A \preccurlyeq B) \to \Box(A \preccurlyeq B)$
9. $\mathbf{A}_2 := (A \prec B) \to \Box(A \prec B)$
10. $\mathbf{WTriv} := (A \preccurlyeq B) \leftrightarrow (\Diamond B \to \Diamond A)$
11. $\mathbf{Triv} := (A \preccurlyeq B) \leftrightarrow (B \to A)$

Some combinations of these axioms are used to obtain Lewis' family of 26 axiomatic systems. However, Lewis does not explore the axiomatic systems obtained from every combination of axioms. For instance, he only ever refers to **4** and **5** as a pair, so he never discusses axiomatic systems like **VT4**. Also, **WTriv** and **Triv** are redundant as they can be shown to be equivalent to a combination of others.

Remark 3.4.2 (Axiomatic Systems). Here are some of the basic axiomatic systems that Lewis describes using the axioms in definition 3.4.1:[4]

[3]Lewis does not name **4**, **5**, $\mathbf{A}_1$, $\mathbf{A}_2$, **WTriv**, and **Triv**. The first two are commonly known in modal logic as **4** and **5**.

[4]Lewis uses **U** to refer to **45** and **A** to refer to $\mathbf{A}_1\mathbf{A}_2$. We won't do this to avoid introducing confusion, since there are no axioms **U** or **A**.

1. **VN**
2. **VT**
3. **VW**
4. **VC**
5. **VS**
6. **V45**
7. $\mathbf{VA_1A_2}$
8. **V45T**
9. $\mathbf{VWA_1A_2}$
10. $\mathbf{VCA_1A_2}$

The following set of theorems show that some axioms imply other axioms.

Theorem 3.4.3.
The following are true:

1. $\vdash_V \Box A \to (\diamond\!\!\!\cdot\, A \lor \neg(\top \prec \bot))$.
2. $\vdash_V \boxdot A \to \diamond\!\!\!\cdot\, A$.
3. $\vdash_{VC} \boldsymbol{N}$.
4. $\vdash_{VN} (\Box A \lor \boxdot A) \to \diamond\!\!\!\cdot\, A$.
5. $\vdash_{VC} \boldsymbol{W}$.
6. $\vdash_{VW} \boldsymbol{T}$.
7. $\vdash_{VT} \boldsymbol{N}$.
8. $\vdash_{VA_1} \boldsymbol{4}$.
9. $\vdash_{VA_2} \boldsymbol{5}$.

Proof. We prove prove two of our theorems here and to avoid clutter we move the remaining proofs to theorem 8.2.1.

1. We provide a derivation for $\vdash_{\mathbf{V}} \Box A \rightarrow (\Diamond A \vee \neg(\top \prec \bot))$.

1.	$\vdash_{\mathbf{V}} \Box A \rightarrow$ $([(A \preccurlyeq \top) \vee (\neg A \preccurlyeq \top)] \rightarrow$ $[\Box A \wedge [(A \preccurlyeq \top) \vee (\neg A \preccurlyeq \top)]])$	Taut
2.	$\vdash_{\mathbf{V}} [\Box A \wedge [(A \preccurlyeq \top) \vee (\neg A \preccurlyeq \top)]] \rightarrow$ $[(\Box A \wedge (A \preccurlyeq \top)) \vee (\Box A \wedge (\neg A \preccurlyeq \top))]$	Taut
3.	$\top \rightarrow (A \vee \neg A)$	Taut
4.	$[(A \preccurlyeq \top) \vee (\neg A \preccurlyeq \top)]$	Cp,3
5.	$\vdash_{\mathbf{V}} \Box A \rightarrow$ $[(\Box A \wedge (A \preccurlyeq \top)) \vee (\Box A \wedge (\neg A \preccurlyeq \top))]$	Pl,1,2,4
6.	$\vdash_{\mathbf{V}} \Box A \rightarrow$ $[(\Box A \wedge (A \preccurlyeq \top)) \vee ((\bot \preccurlyeq \neg A) \wedge (\neg A \preccurlyeq \top))]$	Defn of $\Box$
7.	$\vdash_{\mathbf{V}} ((\bot \preccurlyeq \neg A) \wedge (\neg A \preccurlyeq \top)) \rightarrow (\bot \preccurlyeq \top))$	Trans
8.	$\vdash_{\mathbf{V}} ((\bot \preccurlyeq \neg A) \wedge (\neg A \preccurlyeq \top)) \rightarrow \neg(\top \prec \bot))$	Defn of $\prec$
9.	$\vdash_{\mathbf{V}} \Box A \rightarrow [(A \preccurlyeq \top) \vee \neg(\top \prec \bot)]$	Pl,6,8
10.	$\vdash_{\mathbf{V}} \Box A \rightarrow [\Diamond A \vee \neg(\top \prec \bot)]$	Defn of $\Diamond$

Classical tautologies used:

$$p \rightarrow [q \rightarrow (p \wedge q)]$$
$$[p \wedge (q \vee r)] \rightarrow [(p \wedge q) \vee (p \wedge r)]$$
$$\top \rightarrow (p \vee \neg p).$$

7. We provide a derivation for $\vdash_{\mathbf{VT}} \mathbf{N}$.

1.	$\vdash_{\mathbf{VT}} \Box \bot \rightarrow \bot$	**T**
2.	$\vdash_{\mathbf{VT}} (\bot \preccurlyeq \neg\bot) \rightarrow \bot$	Defn of $\Box$
3.	$\vdash_{\mathbf{VT}} \neg(\neg\bot \prec \bot) \rightarrow \bot$	Defn of $\prec$
4.	$\vdash_{\mathbf{VT}} \top \prec \bot$	Pl,3

Since every instance of **N** is of the form $\top \prec \bot$, then $\vdash_{\mathbf{VT}} \mathbf{N}$.

□

These results tell us that the axiomatic system **VC** is stronger than the axiomatic system **VW** which is stronger than **VT** which is stronger than **VN**. Similarly, we get that $\mathbf{VA_1A_2}$ is stronger than **V45**.

Some rules only hold given certain axioms. For instance, given the axiom, **N**, we can obtain an *inner version* of our *Necessitation* rule.

Theorem 3.4.4 (The Rule INEC)**.**
*Given the axiomatic system **VN**, for any sentence A,*

$$\text{if } \vdash_{VN} A, \text{ then } \vdash_{VN} \boxdot A.$$

Proof.

1.	$\vdash_{\mathbf{VN}} A$	Assumption
2.	$\vdash_{\mathbf{VN}} \neg A \to \bot$	PL,1
3.	$\vdash_{\mathbf{VN}} \bot \preccurlyeq \neg A$	CP,2
4.	$\vdash_{\mathbf{VN}} \top \prec \bot$	**N**
5.	$\vdash_{\mathbf{VN}} (\top \prec \bot) \wedge (\bot \preccurlyeq \neg A)$	PL,3,4
6.	$\vdash_{\mathbf{VN}} ((\top \prec \bot) \wedge (\bot \preccurlyeq \neg A)) \to (\top \prec \neg A)$	By 3.3.2(5)
7.	$\vdash_{\mathbf{VN}} \top \prec \neg A$	**Mp**,6,7
8.	$\vdash_{\mathbf{VN}} \boxdot A$	Defn of $\boxdot$

□

This rule is annotated with INEC and a line number of the sentence it's being applied to.

Additionally, given the axiom, **N**, we can also obtain a version of the Kripke axiom in terms of our inner modality. First we'll prove a lemma.

Lemma 3.4.5.

$$\vdash_{\Sigma} (C \prec (A \vee B)) \to ((C \prec A) \wedge (C \prec B))$$

Proof.

1.	$\vdash_\Sigma A \to (A \lor B)$	Taut
2.	$\vdash_\Sigma (A \lor B) \preccurlyeq A$	Cp,1
3.	$\vdash_\Sigma [(C \prec (A \lor B)) \land ((A \lor B) \preccurlyeq A)] \to (C \prec A)$	By 3.3.2(5)
4.	$\vdash_\Sigma [(A \lor B) \preccurlyeq A] \to [(C \prec (A \lor B)) \to (C \prec A)]$	Pl,3
5.	$\vdash_\Sigma (C \prec (A \lor B)) \to (C \prec A)$	Mp, 2,4
6.	$\vdash_\Sigma B \to (A \lor B)$	Taut
7.	$\vdash_\Sigma (A \lor B) \preccurlyeq B$	Cp,1
8.	$\vdash_\Sigma [(C \prec (A \lor B)) \land ((A \lor B) \preccurlyeq B)] \to (C \prec B)$	By 3.3.2(5)
9.	$\vdash_\Sigma [(A \lor B) \preccurlyeq B] \to [(C \prec (A \lor B)) \to (C \prec B)]$	Pl,8
10.	$\vdash_\Sigma (C \prec (A \lor B)) \to (C \prec B)$	Mp, 7,9
11.	$\vdash_\Sigma (C \prec (A \lor B)) \to [(C \prec A) \land (C \prec B)]$	Pl, 5,10

□

Theorem 3.4.6 (Inner Kripke Axiom)**.**

$$\vdash_{VN} \boxdot(A \land B) \to (\boxdot A \land \boxdot B)$$

Proof.

1.	$\vdash_{\mathbf{VN}} (\neg A \lor \neg B) \leftrightarrow \neg(A \land B)$	Taut
2.	$\vdash_{\mathbf{VN}} [\top \prec (\neg A \lor \neg B)] \to ([\top \prec \neg A] \land [\top \prec \neg B])$	By 3.4.5
3.	$\vdash_{\mathbf{VN}} [\top \prec \neg(A \land B)] \to ([\top \prec \neg A] \land [\top \prec \neg B])$	Eq,1,2
4.	$\vdash_{\mathbf{VN}} \boxdot(A \land B) \to (\boxdot A \land \boxdot B)$	Defn of $\boxdot$

□

3.5 Derivations from a Set of Sentences

We would also like to talk about derivation from a set of sentences. It's important to note that the set of sentences could be infinite in either case. There are two different kinds of derivation from a set and we present both. The first form will be called local derivation from a set.

Definition 3.5.1 (Local Derivation)**.**
Given an axiomatic system, Σ, a set of sentences Γ, and a sentence A, we say Γ

locally derives A in Σ if and only if there exist a finite subset $\{ A_1, \dots, A_n \} \subseteq \Gamma$, such that we have the following theorem in Σ.

$$\vdash_\Sigma (A_1 \wedge \cdots \wedge A_n) \to A.$$

Write:

$$\Gamma \vdash^\ell_\Sigma A.$$

In local derivations, the sentences in Γ do not function like our axioms in Σ. We are more limited in what we can do with them. Ultimately, what $\Gamma \vdash^\ell_\Sigma A$ tells us is that A can be derived from a subset of Γ using modus ponens.

The second form of derivation from a set is called global derivation and it is much stronger than local derivation. Essentially we treat the sentences in our set as if they were axioms in Σ, however, Σ must always be finite while our set could be infinite.

Definition 3.5.2 (Global Derivation)**.**
Given an axiomatic system, Σ, a set of sentences Γ, and a sentence A, we say Γ globally derives A if there exists a finite set of sentences $\{ A_1, \dots, A_n \}$, such that $A_n = A$ and for all A_i, either A_i is an instance of a sentence in Γ, transitivity, strong connectivity, or of a sentence in Σ; or A_i all follow from previous sentences using Mp or Cp.

Write:

$$\Gamma \vdash^g_\Sigma A.$$

Local derivation is stronger than global derivation, in the sense that if we can locally derive something then we can also globally derive it.

Theorem 3.5.3.
Given an axiomatic system, Σ, a set of sentences, Γ, and a sentence A,

$$\text{if } \Gamma \vdash^\ell_\Sigma A \text{ then } \Gamma \vdash^g_\Sigma A.$$

Proof. Let Σ be an axiomatic system, Γ be a set of sentences, and A be a sentence.

Suppose $\Gamma \vdash^\ell_\Sigma A$. Then there exists a finite $\{ A_1, \dots, A_n \} \subseteq \Gamma$ such that $\vdash_\Sigma (A_1 \wedge \cdots \wedge A_n) \to A$. By definition, that means that we have a finite global derivation yielding $\Gamma \vdash_\Sigma (A_1 \wedge \cdots \wedge A_n) \to A$. Additionally, since each A_i in Γ can be introduced as a sentence in a global derivation, then we have a finite global derivation yielding $\Gamma \vdash^\ell_\Sigma (A_1 \wedge \cdots \wedge A_n)$. Together, with rule PL, we can obtain a finite global derivation yielding $\Gamma \vdash^\ell_\Sigma A$, concluding our proof. □

Just as with local entailment in the previous chapter theorem 2.6.7, local derivation also has a deduction theorem.

Theorem 3.5.4 (Deduction Theorem (Syntax version))**.**
Given an axiomatic system, Σ, a set of sentences Γ, and sentences A and B,

$$\Gamma \cup \{ A \} \vdash^\ell_\Sigma B \quad \text{iff} \quad \Gamma \vdash^\ell_\Sigma A \to B.$$

Proof. Let Σ be an axiomatic system, Γ be a set of sentences, and A and B be sentences.

We consider two cases:

Case 1: If $\Gamma = \varnothing$, the definition of $A \vdash^\ell_\Sigma B$ (definition 3.5.1) gives us

$$A \vdash^\ell_\Sigma B \quad \text{iff} \quad \vdash_\Sigma A \to B.$$

Case 2: If $\Gamma \neq \varnothing$, then we do each direction separately.

Suppose $\Gamma \cup \{ A \} \vdash^\ell_\Sigma B$. Then, by definition 3.5.1, there is a finite subset $\{ A_1, \ldots, A_n \} \subseteq \Gamma$ such that we have the theorem $\vdash_\Sigma (A_1 \wedge \cdots \wedge A_n \wedge A) \to B$. Then, by PL we have the theorem $\vdash_\Sigma (A_1 \wedge \cdots \wedge A_n) \to (A \to B)$. Since $\{ A_1, \ldots, A_n \} \subseteq \Gamma$ that means that by definition 3.5.1, $\Gamma \vdash^\ell_\Sigma A \to B$.

Suppose $\Gamma \vdash^\ell_\Sigma A \to B$. Then, by definition 3.5.1, there is a finite subset $\{ A_1, \ldots, A_n \} \subseteq \Gamma$ such that we have the theorem $\vdash_\Sigma (A_1 \wedge \cdots \wedge A_n) \to (A \to B)$. Then, by PL, we have the theorem $\vdash_\Sigma (A_1 \wedge \cdots \wedge A_n \wedge A) \to B$. Finally, since $\{ A_1, \ldots, A_n \} \subseteq \Gamma$ that means that by definition 3.5.1, $\Gamma \cup \{ A \} \vdash^\ell_\Sigma B$. □

3.6 Consistent Sets of Sentences

Analogous to the notion of a satisfiable set of sentences with respect to a frame, we have a notion of a consistent set of sentences with respect to an axiomatic system.

Definition 3.6.1 (Σ-Consistent Set of Sentences)**.**
Given an axiomatic system, Σ, a set of sentences, Γ, is **Σ-consistent** iff $\Gamma \nvdash^{\ell}_{\Sigma} \bot$, otherwise it is **Σ-inconsistent**.

Recall that $\Gamma \nvdash^{\ell}_{\Sigma} \bot$ iff there does not exist a *finite* subset $\left\{ A_1, \dots, A_n \right\} \subseteq \Gamma$ such that $\vdash_{\Sigma} (A_1 \land \cdots \land A_n) \to \bot$.

Essentially, a set of sentences is Σ-consistent if a finite subset of them cannot be used to deduce a contradiction.

We'll present a few useful results about these.

Corollary 3.6.2.
If Γ is Σ-consistent, then any subset of Γ must also be Σ-consistent.

Proof. Follows trivially from the definition. □

Corollary 3.6.3.
If Γ is Σ-inconsistent, then any superset of Γ must also be Σ-inconsistent.

Proof. Follows trivially from the definition. □

Theorem 3.6.4.
Given an axiomatic system, Σ, and a set of sentences, Γ, Γ is Σ-consistent iff there exists a sentence, A, such that $\Gamma \nvdash^{\ell}_{\Sigma} A$.

Proof. Let Σ be an axiomatic system and Γ be a set of sentences. We'll prove each direction.

Suppose Γ is Σ-consistent. Then, by definition 3.6.1, there exists a sentence, $\bot$, such that $\Gamma \nvdash^{\ell}_{\Sigma} \bot$.

Suppose Γ is Σ-inconsistent. Then $\Gamma \vdash^{\ell}_{\Sigma} \bot$, meaning that we have a finite $\left\{ A_1, \dots, A_n \right\} \subseteq \Gamma$ such that $\vdash_{\Sigma} (A_1 \land \cdots \land A_n) \to \bot$. However, since it is the case that for all sentences A, $\vdash_{\Sigma} \bot \to A$ is a theorem, then we have $\vdash_{\Sigma} (A_1 \land \cdots \land A_n) \to A$ by rule Pl. Therefore, there does not exist a sentence, A such that $\Gamma \nvdash^{\ell}_{\Sigma} A$. □

Theorem 3.6.5.
Given an axiomatic system, Σ, a set of sentences, Γ, and a sentence A, $\Gamma \vdash^{\ell}_{\Sigma} A$ iff $\Gamma \cup \{ \neg A \}$ is Σ-inconsistent.

Proof. Let Σ be an axiomatic system, Γ be a set of sentences, and A be a sentence.

For the first direction, suppose $\Gamma \vdash^{\ell}_{\Sigma} A$. Then there is a finite subset $\{ A_1, \dots, A_n \} \subseteq \Gamma$ such that $\vdash_{\Sigma} (A_1 \wedge \cdots \wedge A_n) \to A$. However, since $\vdash_{\Sigma} (A_1 \wedge \cdots \wedge A_n \wedge \neg A) \to \neg A$, then $\vdash_{\Sigma} (A_1 \wedge \cdots \wedge A_n \wedge \neg A) \to \bot$, letting us conclude $\Gamma \cup \{ \neg A \}$ is Σ-inconsistent.

For the second direction, suppose $\Gamma \cup \{ \neg A \}$ is Σ-inconsistent. Then compute:

$$\Gamma \cup \{ \neg A \} \vdash^{\ell}_{\Sigma} \bot$$

iff $\Gamma \vdash^{\ell}_{\Sigma} \neg A \to \bot$

(by theorem 3.5.4)

iff $\Gamma \vdash^{\ell}_{\Sigma} \neg\neg A$

(by definition 2.1.6)

iff $\Gamma \vdash^{\ell}_{\Sigma} A$.

□

Theorem 3.6.6.
Given an axiomatic system, Σ, a set of sentences, Γ, and a sentence A, if $\Gamma \cup \{ A \}$ and $\Gamma \cup \{ \neg A \}$ are both Σ-inconsistent, then Γ is Σ-inconsistent.

Proof. Let Σ be an axiomatic system, Γ be a set of sentences, and A be a sentence.

Suppose $\Gamma \cup \{ A \}$ and $\Gamma \cup \{ \neg A \}$ are both Σ-inconsistent. Then $\Gamma \vdash^{\ell}_{\Sigma} \neg A$ and $\Gamma \vdash^{\ell}_{\Sigma} A$ by theorem 3.6.5. However, then there must exist a finite subset of Γ such that $\Gamma \vdash^{\ell}_{\Sigma} A \wedge \neg A$. Therefore $\Gamma \vdash^{\ell}_{\Sigma} \bot$, and Γ is Σ-inconsistent. □

3.7 Lindenbaum's Lemma

Since we can add sentences to a Σ-consistent set, it makes sense to ask if one can have a Σ-consistent set that contains as many sentences as possible. It is, we call these maximally Σ-consistent sets.

Definition 3.7.1 (Maximally Σ-Consistent Set of Sentences)**.**
Given an axiomatic system, Σ, Σ-consistent set of sentences, Γ, is maximally Σ-consistent iff for every sentence, A, Γ is Σ-consistent and either $A \in \Gamma$ or $\neg A \in \Gamma$.

An important classical result is that given any Σ-consistent set it is possible to add sentences to it in order to obtain a maximally Σ-consistent set. We'll often refer to this as extending a Σ-consistent set into a maximally Σ-consistent set. We'll refer to this result a lot in coming chapters.

Theorem 3.7.2 (Lindenbaum's Lemma)**.**
Any Σ-consistent set of sentences can be extended to a maximal Σ-consistent set.

This is a standard proof.

Proof. Suppose Γ is a Σ-consistent set of sentences. First, recall that our language, L_V is countable by theorem 2.1.5 and therefore its sentences have an enumeration. Assume such an ordering and let $L_V = (A_0, A_1, \dots)$.

This lets us obtain the recursive definition. Let $\Gamma_0 = \Gamma$, and

$$\Gamma_{n+1} := \begin{cases} \Gamma_n \cup \{ A_n \} & \text{if } \Gamma_n \cup \{ A_n \} \text{ is } \Sigma\text{-consistent} \\ \Gamma_n \cup \{ \neg A_n \} & \text{otherwise.} \end{cases}$$

Next, by induction on n we prove that $\forall n \in \mathbb{N}, \Gamma_n$ is Σ-consistent.

For the base case, note that Γ_0 is Σ-consistent by assumption. Let $k \in \mathbb{N}$ be arbitrary. Suppose Γ_k is Σ-consistent.

Either $\Gamma_k \cup \{ A_k \}$ is Σ-consistent or else it is not. In the first case, Γ_{k+1} is Σ-consistent by definition. On the other hand, if $\Gamma_k \cup \{ A_k \}$ is Σ-inconsistent then $\Gamma_k \cup \{ \neg A_k \}$ must be Σ-consistent (else Γ_k itself was Σ-inconsistent by theorem 3.6.6). But in this case, $\Gamma_{k+1} = \Gamma_k \cup \{ \neg A_k \}$ is Σ-consistent.

This completes our induction, and we have that $\forall n \in \mathbb{N}, \Gamma_n$.

Next define

$$\Gamma' = \bigcup_{n \in \mathbb{N}} \Gamma_n.$$

If Γ' were Σ-inconsistent, then there would exist a finite set $\{ A_1, \dots, A_n \} \subseteq \Gamma'$ such that $\vdash_\Sigma (A_1 \wedge \cdots \wedge A_n) \to \bot$ but this would mean that $\{ A_1, \dots, A_n \}$ would itself be a finite Σ-inconsistent subset of Γ'. However, such a finite set would be a subset of a Σ-consistent Γ_n (for some n). By corollary 3.6.2, this gives us a contradiction telling us that Γ' must be Σ-consistent. Moreover, given that Γ' is Σ-consistent it trivially satisfies our definition of maximally Σ-consistent. This completes the proof. □

It will simplify our discussion if we introduce some notation for maximally Σ-consistent sets.

Definition 3.7.3.
Given an axiomatic system, Σ, we denote the set of maximally Σ-consistent sets by $\mathscr{W}_\Sigma$.

This theorem gives us some properties for maximally Σ-consistent sets.

Theorem 3.7.4.
Given an axiomatic system, Σ, and a $\Delta \in \mathscr{W}_\Sigma$, the following hold:

1. *Δ is deductively closed in Σ. That is to say:*

   For all sentences, A, if $\Delta \vdash^{\ell}_{\Sigma} A$ then $A \in \Delta$.

2. $\Sigma \subseteq \Delta$.

3. $\perp \notin \Delta$.

4. *For all sentences A and B, $(A \to B) \in \Delta$ iff $A \notin \Delta$ or $B \in \Delta$.*

Proof. Let Σ be an axiomatic system and let $\Delta \in \mathscr{W}_\Sigma$. We prove each statement.

1. Let A be an arbitrary sentence. Suppose $\Delta \vdash^{\ell}_{\Sigma} A$. Suppose for the sake of contradiction that $A \notin \Delta$. Then, since Δ is maximally Σ-consistent we must have $\neg A \in \Delta$, but then $\Delta \vdash^{\ell}_{\Sigma} \neg A$ (by definition 3.5.1, since $\vdash_\Sigma \neg A \to \neg A$). Together with our assumption this gives us $\Delta \vdash^{\ell}_{\Sigma} \perp$ contradicting the fact that Δ is maximally Σ-consistent. Therefore $A \in \Delta$.

2. Let $A \in \Sigma$ be arbitrary, then we have $\vdash_\Sigma A$. This implies that $\Delta \vdash^{\ell}_{\Sigma} A$ and since Δ is deductively closed theorem 3.7.4(item 1) we have that $A \in \Delta$. Since $A \in \Sigma$ was arbitrary, then $\Sigma \subseteq \Delta$.

3. Suppose for the sake of contradiction that $\bot \in \Delta$, then $\Delta \vdash^{\ell}_{\Sigma} \bot$ which contradicts the definition of Δ being Σ-consistent. Therefore $\bot \notin \Delta$.

4. Let A and B be arbitrary sentences. Suppose $A \notin \Delta$ or $B \in \Delta$. We have two cases, in the first case, if $A \notin \Delta$ then $\neg A \in \Delta$ and therefore $(A \to B) \in \Delta$. In the second case, if $B \in \Delta$ then $(A \to B) \in \Delta$.

□

Chapter 4

Soundness

We've now described a syntax and a semantics for our logic and we'd like to start showing how the two correspond. In this chapter we'll prove that our semantics is "sound" with respect to our logic:

> "If a sentence is derivable in the syntax then it is valid in the semantics."

Taking the contrapositive of this statement we may interpret it as saying:

> "No invalid sentence may be derived."

We'll prove our soundness theorems by induction over the length of derivations. For our base cases we'll need that our axioms, Trans and Connex, are valid. Then, for our inductive step, we'll need to show that our rules, Mp and Cp, preserve validity. We'll prove lemmas for each of these over the first couple sections of this chapter. We also have a few different versions of soundness and so we'll have a few different versions of our lemmas.

Our versions of soundness are as follows. Given an axiomatic system $\Sigma = \{ A_1, A_2, \dots, A_n \}$ and the class of frames obtained from the intersection of a family of classes of frames, $\mathscr{F} = \bigcap \{ \mathscr{F}_1, \mathscr{F}_2, \dots, \mathscr{F}_n \}$, such that $\vDash_{\mathscr{F}_1} A_1$, $\vDash_{\mathscr{F}_2} A_2$, …, $\vDash_{\mathscr{F}_n} A_n$, then for every sentence A, and set of sentences Γ,

1. Soundness: If $\vdash_\Sigma A$ then $\vDash_{\mathscr{F}} A$.

2. Strong (Local) Soundness: If $\Gamma \vdash^{\ell}_{\Sigma} A$ then $\Gamma \vDash^{\ell}_{\mathscr{F}} A$.

3. Strong (Local) Soundness (Alternate Form): If Γ is $\mathscr{F}$-satisfiable, then Γ is Σ-consistent.

We will first prove (1), then show that (1) implies (2), and finally show that (3) is equivalent to (2).

4.1 Axioms Are Valid

We first prove that all instances of transitivity are valid in every sphere model.

Theorem 4.1.1.
Given sentences A, B, and C, for every sphere model, $M = (W, \mathscr{O}, v)$,

$$[\![((A \preccurlyeq B) \wedge (B \preccurlyeq C)) \rightarrow (A \preccurlyeq C)]\!] = W.$$

Proof. Let A, B, and C be sentences. Let $M = (W, \mathscr{O}, v)$ be an arbitrary sphere model. Let $w \in W$ be arbitrary.

We will prove that $[\![((A \preccurlyeq B) \wedge (B \preccurlyeq C))]\!] \subseteq [\![(A \preccurlyeq C)]\!]$ and use lemma 2.3.6 to obtain the result.

$$
\begin{aligned}
& w \in [\![(A \preccurlyeq B) \wedge (B \preccurlyeq C)]\!] \\
& \quad \text{iff } w \in [\![A \preccurlyeq B]\!] \cap [\![B \preccurlyeq C]\!] \\
& \quad \text{iff } w \in [\![A \preccurlyeq B]\!] \\
& \qquad \text{and } w \in [\![B \preccurlyeq C]\!] \\
& \quad \text{iff } w \in \{\, v \in W \mid \forall S \in \mathscr{O}(v), ([\![B]\!] \bullet S \rightarrow [\![A]\!] \bullet S) \,\} \\
& \qquad \text{and } w \in \{\, v \in W \mid \forall S \in \mathscr{O}(v), ([\![C]\!] \bullet S \rightarrow [\![B]\!] \bullet S) \,\} \\
& \quad \text{iff } \forall S \in \mathscr{O}(w), ([\![B]\!] \bullet S \rightarrow [\![A]\!] \bullet S) \\
& \qquad \text{and } \forall S \in \mathscr{O}(w), ([\![C]\!] \bullet S \rightarrow [\![B]\!] \bullet S) \\
& \quad \text{iff } \forall S \in \mathscr{O}(w), ([\![B]\!] \bullet S \rightarrow [\![A]\!] \bullet S) \wedge ([\![C]\!] \bullet S \rightarrow [\![B]\!] \bullet S) \\
& \quad \text{only if } \forall S \in \mathscr{O}(w), ([\![C]\!] \bullet S \rightarrow [\![A]\!] \bullet S) \\
& \quad \text{iff } w \in \{\, v \in W \mid \forall S \in \mathscr{O}(v), ([\![C]\!] \bullet S \rightarrow [\![A]\!] \bullet S) \,\} \\
& \quad \text{iff } w \in [\![A \preccurlyeq C]\!].
\end{aligned}
$$

Therefore, $[\![((A \preccurlyeq B) \wedge (B \preccurlyeq C))]\!] \subseteq [\![(A \preccurlyeq C)]\!]$. By applying the lemma 2.3.6 we obtain $[\![((A \preccurlyeq B) \wedge (B \preccurlyeq C)) \rightarrow (A \preccurlyeq C)]\!] = W$. □

Finally we show that all instances of strong connectivity are valid in every sphere model.

Theorem 4.1.2.
Given sentences A and B, for every sphere model, $M = (W, \mathcal{O}, v)$,

$$[\![(A \preccurlyeq B) \vee (B \preccurlyeq A)]\!] = W.$$

Proof. Let A and B be sentences. Let $M = (W, \mathcal{O}, v)$ be an arbitrary sphere model.

Suppose for the sake of contradiction that $[\![(A \preccurlyeq B) \vee (B \preccurlyeq A)]\!] \neq W$. That is to say, suppose there exists a world, $w \in W$, such that $w \notin [\![(A \preccurlyeq B) \vee (B \preccurlyeq A)]\!]$ and proceed:

$$\begin{aligned}
&w \notin [\![(A \preccurlyeq B) \vee (B \preccurlyeq A)]\!] \\
&\quad \text{iff } w \notin ([\![A \preccurlyeq B]\!] \cup [\![B \preccurlyeq A]\!]) \\
&\quad \text{iff } w \notin [\![A \preccurlyeq B]\!] \\
&\quad\quad \text{and } w \notin [\![B \preccurlyeq A]\!] \\
&\quad \text{iff } w \in [\![\neg(A \preccurlyeq B)]\!] \\
&\quad\quad \text{and } w \in [\![\neg(B \preccurlyeq A)]\!] \\
&\quad \text{iff } w \in [\![B \prec A]\!] \\
&\quad\quad \text{and } w \in [\![A \prec B]\!] \\
&\quad \text{iff } w \in \{\, v \in W \mid \exists S \in \mathcal{O}(v), ([\![B]\!] \bullet S \wedge [\![A]\!] \not\bullet S) \,\} \\
&\quad\quad \text{and } w \in \{\, v \in W \mid \exists S \in \mathcal{O}(v), ([\![A]\!] \bullet S \wedge [\![B]\!] \not\bullet S) \,\} \\
&\quad \text{iff } \exists S \in \mathcal{O}(w), ([\![B]\!] \bullet S \wedge [\![A]\!] \not\bullet S) \\
&\quad\quad \text{and } \exists S \in \mathcal{O}(w), ([\![A]\!] \bullet S \wedge [\![B]\!] \not\bullet S).
\end{aligned}$$

Let $S_1, S_2 \in \mathcal{O}(w)$ be spheres satisfying each of our statements. Then, we have

$$([\![B]\!] \bullet S_1 \wedge [\![A]\!] \not\bullet S_1) \quad \text{and} \quad ([\![A]\!] \bullet S_2 \wedge [\![B]\!] \not\bullet S_2).$$

However, by definition our spheres are nested, so it must be the case that either $S_1 \subseteq S_2$ or $S_2 \subseteq S_1$. Without loss of generality, suppose $S_1 \subseteq S_2$. Then, it follows that since, $[\![B]\!] \bullet S_1$, it must be the case that $[\![B]\!] \bullet S_2$ but we have deduced from our assumption that $[\![B]\!] \not\bullet S_2$ giving us a contradiction.

Therefore, our initial assumption must be false and we can conclude that, $[\![(A \preccurlyeq B) \vee (B \preccurlyeq A)]\!] = W$, completing the proof. □

4.2 Rules Preserve Validity

We have now shown that all instances of our axioms are valid over all sphere models. Next, we need to show that our rules, Mp and Cp, preserve validity.

First we prove that Mp preserves truth (this is a stronger property than preserving validity).

Lemma 4.2.1.
Given sentences A and B, for every sphere model, $M = (W, \mathcal{O}, v)$, and for every $w \in W$,

$$\text{if } M, w \models A \text{ and } M, w \models (A \rightarrow B) \text{ then } M, w \models B.$$

Proof. Let A and B be sentences. Suppose $M = (W, \mathcal{O}, v)$ is a sphere model and $w \in W$.

Suppose $M, w \models A$ and $M, w \models (A \rightarrow B)$. Then, $w \in [\![A]\!]$ and let $w \in [\![A \rightarrow B]\!]$. By definition of $[\![A \rightarrow B]\!]$ definition 2.3.3(item 3), $w \in (W - [\![A]\!]) \cup [\![B]\!]$. So, we have that $w \in (W - [\![A]\!])$ or $w \in [\![B]\!]$. However, since $w \in [\![A]\!]$ then $w \notin (W - [\![A]\!])$. So $w \in [\![B]\!]$, which gives us that $M, w \models B$ as desired. □

Next we prove that Mp preserves validity.

Theorem 4.2.2.
Given sentences A and B, and a class of frames $\mathcal{F}$, if $\Gamma \models^{\ell}_{\mathcal{F}} A$ and $\Gamma \models^{\ell}_{\mathcal{F}} (A \rightarrow B)$, then $\Gamma \models^{\ell}_{\mathcal{F}} B$.

Proof. Let A and B be sentences. Let $\mathcal{F}$ be an arbitrary class of frames. Suppose $\Gamma \models^{\ell}_{\mathcal{F}} A$ and $\Gamma \models^{\ell}_{\mathcal{F}} (A \rightarrow B)$.

Let $F \in \mathscr{F}$ be an arbitrary frame. Let $M = (W, \mathcal{O}, v)$ be an arbitrary model over F. Let $w \in W$ be arbitrary. Suppose $M, w \vDash \Gamma$.

Then we have $M, w \vDash A$ and $M, w \vDash (A \to B)$. By lemma 4.2.1 it follows that $M, w \vDash B$. Then, since $w \in W$, M over F, and $F \in \mathscr{F}$, were all arbitrary, it follows that $\Gamma \vDash^{\ell}_{\mathscr{F}} B$. □

We also get that MP holds for entailment as well.

Corollary 4.2.3.
Given sentences A and B, and a class of frames $\mathscr{F}$, if $\vDash_{\mathscr{F}} A$ and $\vDash_{\mathscr{F}} (A \to B)$, then $\vDash_{\mathscr{F}} B$.

Proof. The proof follows from theorem 4.2.2 by considering the case when $\Gamma = \varnothing$. □

We also have that MP preserves validity in terms of global entailment.

Corollary 4.2.4.
Given sentences A and B, and a class of frames $\mathscr{F}$, if $\Gamma \vDash^{g}_{\mathscr{F}} A$ and $\Gamma \vDash^{g}_{\mathscr{F}} (A \to B)$, then $\Gamma \vDash^{g}_{\mathscr{F}} B$.

Proof. The proof is a similar argument to the one used to prove theorem 4.2.2. □

Theorem 4.2.5.
For all $n \in \mathbb{N}$, such that $n \geq 1$. Given sentences A, $B_1, \ldots, B_n$ and a class of frames, $\mathscr{F}$, if $\Gamma \vDash^{g}_{\mathscr{F}} A \to (B_1 \vee \cdots \vee B_n)$ then $\Gamma \vDash^{g}_{\mathscr{F}} (B_1 \preccurlyeq A) \vee \cdots \vee (B_n \preccurlyeq A)$.

Proof. Let $n \in \mathbb{N}$, such that $n \geq 1$. Let A, $B_1, \ldots, B_n$ be sentences. Let $\mathscr{F}$ be an arbitrary class of sphere frames.

Suppose $\Gamma \vDash^{g}_{\mathscr{F}} A \to (B_1 \vee \cdots \vee B_n)$. Suppose for the sake of contradiction that $\Gamma \nvDash^{g}_{\mathscr{F}} (B_1 \preccurlyeq A) \vee \cdots \vee (B_n \preccurlyeq A)$.

Then there exists a frame, $F \in \mathscr{F}$, and a model, M, over F, such that $M \vDash \Gamma$ but $M \nvDash (B_1 \preccurlyeq A) \vee \cdots \vee (B_n \preccurlyeq A)$. Therefore, there exists a model $M = (W, \mathcal{O}, v)$ over F, such that $M \vDash \Gamma$ but $M \nvDash (B_1 \preccurlyeq A) \vee \cdots \vee (B_n \preccurlyeq A)$.

So, there exists a $w \in W$, such that $M, w \not\models (B_1 \preccurlyeq A) \vee \cdots \vee (B_n \preccurlyeq A)$. Then compute:

$$\begin{aligned}
&w \notin [\![(B_1 \preccurlyeq A) \vee \cdots \vee (B_n \preccurlyeq A)]\!] \\
&\qquad \text{iff } w \notin \left([\![B_1 \preccurlyeq A]\!] \cup \cdots \cup [\![B_n \preccurlyeq A]\!]\right) \\
&\qquad \text{iff } w \notin [\![B_1 \preccurlyeq A]\!] \text{ and } \cdots \text{ and } w \notin [\![B_n \preccurlyeq A]\!] \\
&\qquad \text{iff } w \in [\![\neg(B_1 \preccurlyeq A)]\!] \text{ and } \cdots \text{ and } w \in [\![\neg(B_n \preccurlyeq A)]\!] \\
&\qquad \text{iff } w \in [\![A \prec B_1]\!] \text{ and } \cdots \text{ and } w \in [\![A \prec B_n]\!].
\end{aligned}$$

So, for each i, such that $1 \le i \le n$, we have $w \in [\![A \prec B_i]\!]$.

By theorem 2.3.5(8), we have

$$\begin{aligned}
&w \in [\![A \prec B_i]\!] \\
&\qquad \text{iff } w \in \left\{ v \in W \;\middle|\; \exists S_i \in \mathcal{O}(v), ([\![A]\!] \bullet S_i \wedge [\![B_i]\!] \not\bullet S_i) \right\} \\
&\qquad \text{iff } \exists S_i \in \mathcal{O}(w), ([\![A]\!] \bullet S_i \wedge [\![B_i]\!] \not\bullet S_i).
\end{aligned}$$

So, there exists n spheres in $\mathcal{O}(w)$ (one for each B_i) such that they all intersect $[\![A]\!]$ but each one fails to intersect its corresponding $[\![B_i]\!]$. Taking the finite intersection,

$$S := \bigcap_{1 \le i \le n} S_i,$$

gives us a sphere such that $[\![A]\!] \bullet S$ and for all i, $[\![B_i]\!] \not\bullet S$.

Next, recall our initial assumption and the fact that $M \models \Gamma$ to obtain $M \models A \to (B_1 \vee \cdots \vee B_n)$. Let $v \in [\![A]\!] \cap S$, it follows that $M, v \models A$ and $M, v \models A \to (B_1 \vee \cdots \vee B_n)$. Then, by lemma 4.2.1, we get $M, v \models (B_1 \vee \cdots \vee B_n)$. This gives us $v \in [\![B_1 \vee \cdots \vee B_n]\!]$, or equivalently $v \in [\![B_1]\!] \cup \cdots \cup [\![B_n]\!]$. However, recalling that $v \in S$, this tells us there must exist some B_i such that $v \in [\![B_i]\!] \cap S$, but this contradicts our earlier result that for all i, $[\![B_i]\!] \not\bullet S$.

Therefore, by contradiction we get that,

$$\Gamma \models^{g}_{\mathcal{F}} (B_1 \preccurlyeq A) \vee \cdots \vee (B_n \preccurlyeq A),$$

finishing our proof. □

Corollary 4.2.6.
For all $n \in \mathbb{N}$, such that $n \geq 1$. Given sentences A, $B_1, \dots, B_n$ and a class of frames, $\mathscr{F}$, if $\vDash^g_{\mathscr{F}} A \to (B_1 \vee \cdots \vee B_n)$ then $\vDash^g_{\mathscr{F}} (B_1 \preccurlyeq A) \vee \cdots \vee (B_n \preccurlyeq A)$.

Proof. Follows from theorem 4.2.5 by considering the case where $\Gamma = \varnothing$. □

4.3 Soundness Theorem

Theorem 4.3.1 (Soundness Theorem)**.**
Given a family of classes of frames, $\{ \mathscr{F}_1, \dots, \mathscr{F}_n \}$, and a set of sentences, $\{ A_1, \dots, A_n \}$. If for all i we have

$$\vDash_{\mathscr{F}_i} A_i$$

then for every sentence, A,

$$\text{if } \vdash_{\Sigma} A \text{ then } \vDash_{\mathscr{F}} A$$

where $\Sigma = \{ A_1, \dots, A_n \}$ and $\mathscr{F} = \bigcap \{ \mathscr{F}_1, \dots, \mathscr{F}_n \}$.

Proof. Let $\mathscr{F}_1, \dots, \mathscr{F}_n$ be a family of classes of frames and let $A_1, \dots, A_n$ be a family of sentences. Suppose that for all i we have

$$\vDash_{\mathscr{F}_i} A_i.$$

Let A be an arbitrary sentence and suppose $\vdash_{\Sigma} A$ where $\Sigma = \{ A_1, \dots, A_n \}$. We wish to show that $\vDash_{\mathscr{F}} A$ where $\mathscr{F} = \bigcap \{ \mathscr{F}_1, \dots, \mathscr{F}_n \}$.

Since $\vdash_{\Sigma} A$, that means there exists a finite derivation of A. We prove $\vDash_{\mathscr{F}} A$ by induction on length of the derivation of $\vdash_{\Sigma} A$.

Base Case:

Suppose $\vdash_{\Sigma} A$ has a derivation of length 1. We have four cases:

1. If A is an instance of an $A_i \in \Sigma$ then $\vDash_{\mathscr{F}_i} A_i$ by our assumption. Then by theorem 2.4.5 we have $\vDash_{\mathscr{F}_i} A$ and therefore $\vDash_{\mathscr{F}} A$.

2. If A is an instance of a classical tautology then by theorem 2.4.7 we know it must be valid over all models.

3. If A is an instance of **Transitivity** then by theorem 4.1.1 it must be valid over all models.

4. If A is an instance of **Strong Connectivity** then by theorem 4.1.2 it must be valid over all models.

Inductive Case: Next we consider the case where the length of the derivation of $\vdash_\Sigma A$ is greater than 1.

Inductive Hypothesis: Let $k \geq 1$ be arbitrary. Suppose that for derivations of $\vdash_\Sigma A$ with length at most k, if $\vdash_\Sigma \ A$ then $\vDash_{\mathscr{F}} A$. We wish to prove that this holds for the case when derivations have length $k+1$.

We have three cases.

1. If A is an instance of an $A_i \in \Sigma$, a classical tautology, transitivity, or strong connectivity, then we proceed as in the base case.

2. If A was obtained by Modus Ponens from sentences B and $B \to A$, then both of those sentences have derivations of length at most k and therefore we have $\vDash_{\mathscr{F}} \ B$ and $\vDash_{\mathscr{F}} \ B \to A$ by inductive hypothesis. Then by corollary 4.2.3, we have that $\vDash_{\mathscr{F}} A$.

3. If A was obtained by Comparative Possibility from a sentence B, then B has a derivation of length at most k and therefore, by the inductive hypothesis, we have $\vDash_{\mathscr{F}} B$. Then by corollary 4.2.6, we have that $\vDash_{\mathscr{F}} A$.

□

Remark 4.3.2. When we consider an empty family of classes of frames and an empty set of sentences we get the special case, if $\vdash_{\mathbf{V}} A$ then $\vDash_{\mathscr{F}} A$ where $\mathscr{F}$ is the class of all frames.

Next, we'd like to show that strong (local) soundness is implied by soundness. For that we'll need this lemma.

Lemma 4.3.3 (Soundness Implies Strong (Local) Soundness)**.**
Given an axiomatic system, Σ and a class of frames $\mathscr{F}$.

If for all sentences A,

$$\vdash_\Sigma A \text{ implies } \vDash_{\mathscr{F}} A,$$

then for all sets of sentences Γ, and sentences A,

$$\Gamma \vdash^{\ell}_{\Sigma} A \text{ implies } \Gamma \vDash^{\ell}_{\mathscr{F}} A.$$

Proof. Let Σ be an axiomatic system, let $\mathscr{F}$ be a class of frames. Suppose that for all sentences A, $\vdash_\Sigma A$ implies $\vDash_\mathscr{F} A$. Let Γ be an arbitrary set of sentences and let A be an arbitrary sentence. Suppose $\Gamma \vdash^\ell_\Sigma A$. We wish to show $\Gamma \vDash^\ell_\mathscr{F} A$.

Then there exists a finite subset $\{ A_1, \dots, A_n \} \subseteq \Gamma$ such that $\vdash_\Sigma (A_1 \wedge \cdots \wedge A_n) \to A$. Therefore, by our assumption, it follows that $\vDash_\mathscr{F} (A_1 \wedge \cdots \wedge A_n) \to A$. By the deduction theorem 2.6.7, $\{ A_1, \dots, A_n \} \vDash^\ell_\mathscr{F} A$. Then by the monotony theorem 2.6.10 we get $\Gamma \vDash^\ell_\mathscr{F} A$. Moreover, since Γ and A were arbitrary, then we have that for all sets of sentences Γ, and sentences A, $\Gamma \vdash^\ell_\Sigma A$ implies $\Gamma \vDash^\ell_\mathscr{F} A$, as desired. □

Now our lemma trivially gives us strong (local) soundness.

Theorem 4.3.4 (Strong (Local) Soundness Theorem)**.**
Given a family of classes of frames, $\{ \mathscr{F}_1, \dots, \mathscr{F}_n \}$*, and a set of sentences,* $\{ A_1, \dots, A_n \}$*. If for all* i *we have*

$$\vDash_{\mathscr{F}_i} A_i$$

then for every sentence, A*, and set of sentences* Γ*,*

$$\text{if } \Gamma \vdash^\ell_\Sigma A \text{ then } \Gamma \vDash^\ell_\mathscr{F} A$$

where $\Sigma = \{ A_1, \dots, A_n \}$ *and* $\mathscr{F} = \bigcap \{ \mathscr{F}_1, \dots, \mathscr{F}_n \}$*.*

Proof. This follows trivially from theorem 4.3.1 and lemma 4.3.3. □

It's also easy to see that strong (local) soundness implies soundness by considering the case where $\Gamma = \varnothing$.

It will at times be useful to restate the strong soundness theorem in terms of $\mathscr{F}$-satisfiability and Σ-consistency. To do so we first make the following simple observation.

Theorem 4.3.5.
Given an axiomatic system Σ*, and a class of frames,* $\mathscr{F}$*, the following are equivalent*

1. $\forall A, \forall \Gamma$*, if* $\Gamma \vdash^\ell_\Sigma A$ *then* $\Gamma \vDash^\ell_\mathscr{F} A$*,*

2. $\forall \Gamma$*, if* Γ *is* $\mathscr{F}$*-satisfiable then* Γ *is* Σ*-consistent.*

Proof. Let Σ, be an axiomatic system, and let $\mathscr{F}$, be a class of frames.

We prove each direction separately.

First, suppose (1) is true, will we prove (2). Let Γ be arbitrary, suppose Γ is $\mathscr{F}$-satisfiable. Then it follows that:

$$\begin{aligned} &\Gamma \text{ is } \mathscr{F}\text{-satisfiable} \\ &\quad \text{iff } \Gamma \not\vDash^{\ell}_{\mathscr{F}} \bot \\ &\qquad \text{(by theorem 2.7.2)} \\ &\quad \text{only if } \Gamma \nvdash^{\ell}_{\Sigma} \bot \\ &\qquad \text{(by (1))} \\ &\quad \text{iff } \Gamma \text{ is } \Sigma\text{-consistent.} \end{aligned}$$

Since Γ was arbitrary then we've proven that (1) implies (2).

Next we prove the other direction. Suppose (2) is true, we will prove the contrapositive of (1). Let A be arbitrary, let Γ be arbitrary. Then it follows that:

$$\begin{aligned} &\Gamma \not\vDash^{\ell}_{\mathscr{F}} A \\ &\quad \text{iff } \Gamma \cup \{ \neg A \} \text{ is } \mathscr{F}\text{-satisfiable} \\ &\qquad \text{(by theorem 2.7.3)} \\ &\quad \text{only if } \Gamma \cup \{ \neg A \} \text{ is } \Sigma\text{-consistent} \\ &\qquad \text{(by (2))} \\ &\quad \text{iff } \Gamma \nvdash^{\ell}_{\Sigma} A \\ &\qquad \text{(by theorem 3.6.5).} \end{aligned}$$ □

Hence if $\Gamma \not\vDash^{\ell}_{\mathscr{F}} A$ then $\Gamma \nvdash^{\ell}_{\Sigma} A$. Take the contrapositive of this statement. Next, since A and Γ were arbitrary then we've proven that (2) implies (1).

This concludes the proof.

The restated version of the strong (local) soundness theorem is as follows.

Corollary 4.3.6 (Strong (Local) Soundness Theorem (Alternate Form))**.** *Given a family of classes of frames,* $\{ \mathscr{F}_1, \dots, \mathscr{F}_n \}$*, and a set of sentences,* $\{ A_1, \dots, A_n \}$*.*

If for all i we have

$$\vDash_{\mathscr{F}_i} A_i$$

then for every set of sentences Γ,

$$\text{if } \Gamma \text{ is } \mathscr{F}\text{-satisfiable then } \Gamma \text{ is } \Sigma\text{-consistent}.$$

where $\Sigma = \{ A_1, \ldots, A_n \}$ and $\mathscr{F} = \bigcap \{ \mathscr{F}_1, \ldots, \mathscr{F}_n \}$.

Proof. Simply apply theorem 4.3.5 and observe we obtain the statement of theorem 4.3.4. □

Chapter 5

Canonical Models

In the last model we showed that our semantics is sound with respect to our syntax. The next thing we'd like to show is that our semantics is "complete" with respect to our syntax. That means that "if a sentence is valid in our semantics, then it can be derived in our syntax," or equivalently, that any valid sentence over a class of frames (characterized by a set of axioms) must also be derivable in the axiomatic system given by those axioms. Our strategy will be to prove the contrapositive of this statement — that is, we will prove that if a sentence isn't derivable then there must exist a model (over the frame in the class of frames in question) where the sentence is invalid. At first glance this may sound like we'd have to deduce a general strategy for constructing countermodels, but fortunately we can achieve this by constructing one single countermodel for our axiomatic system, called the *Canonical Model*. Additionally we'll need to prove that said Canonical Model is actually in the class of frames we're interested in.

How does a canonical model work? Well, first let's back up a little bit. Ordinarily, a sentence will be valid over a model if and only if it is true at all the worlds in the model. Moreover, given a model we can consider the set of sentences that are valid over the model. This set characterizes the model in the sense that a sentence is invalid over the model if and only if the sentence is not contained in the model's set of valid sentences. With this key insight we can see how the canonical model works. Sets of sentences like this will act as our worlds in our canonical model and this way, a sentence will only be valid in the canonical model if it is valid in every model. Of course, defining the canonical model so that it is an actual sphere model is non-trivial and we will dedicate the entirety of this chapter to this.

To construct a canonical model we will want to define the spheres around a world, Δ (I'll use Δ here instead of w as a reminder that the worlds in the canonical model are maximally Σ-consistent sets of sentences). The way to do this is a little convoluted. We'll first define a set of objects called "cuts around Δ." Each cut around Δ will be an inconsistent set of sentences. The idea here is that since cuts contain all the inconsistent stuff we "don't want" then by doing something analogous to taking a complement of a cut we will obtain a corresponding "co-sphere." Later, we'll obtain spheres by taking arbitrary unions of "co-spheres" and this finally gets us our canonical model. At this point we still need to prove that the result is indeed a sphere model and to do that we need to show that cuts, co-spheres, and spheres each satisfy a number of properties. Finally we can prove the truth lemma and that gives us everything we need in order to start proving the completeness theorem in the next chapter.

Lewis' construction of canonical models relies on a notion of a cut and a co-sphere which we need to define first. I provide Lewis' definitions for these notions and some slightly simpler definitions with proofs that they are equivalent.

5.1 Cuts and Co-Spheres

First, we will define what a cut is. Note that for each $\Delta \in \mathscr{W}_\Sigma$ we may obtain many cuts around Δ.

Definition 5.1.1 (A Cut Around $\Delta \in \mathscr{W}_\Sigma$)**.**
Let Σ be an axiomatic system. Given a $\Delta \in \mathscr{W}_\Sigma$, a set of sentences, Ψ_Δ, is called a **cut around** Δ iff it satisfies the following two properties:

1. $\bot \in \Psi_\Delta$, and

2. If $B \in \Psi_\Delta$ and $A \notin \Psi_\Delta$, then $(A \prec B) \in \Delta$.

There may be many cuts around Δ. We may define the collection of all such cuts as

$$\textsc{Cuts}(\Delta) := \left\{ \Psi_\Delta \;\middle|\; \Psi_\Delta \text{ is a cut around } \Delta \right\}.$$

Remark 5.1.2. Lewis (2001, Section 6.1 p. 127) provides a slightly different definition of a cut where item 2 of definition 5.1.1 is stated in the following equivalent way.

2. If $B \in \Psi_\Delta$ and $A \notin \Psi_\Delta$, then $(B \preccurlyeq A) \notin \Delta$.

Equivalence follows from the fact that Δ is maximally Σ-consistent. Since Δ is maximally Σ-consistent then $\neg(B \preccurlyeq A) \in \Delta$ which is equivalent to $(A \prec B) \in \Delta$.

The idea is that each cut will contain an inconsistent set of "bad" sentences in such a way that by considering the negated, "good," sentences we'll later be able to obtain spheres. It will be helpful to show some properties about how cuts behave before continuing on. First we will prove that the set of all sentences is itself a cut. This, the largest cut, is in some sense a degenerate example because it includes all sentences as "bad" sentences.

Theorem 5.1.3.
Let Σ be an axiomatic system. Given a $\Delta \in \mathscr{W}_\Sigma$, if $\Psi_\Delta = L_V$ then Ψ_Δ is a cut around Δ.

Proof. Let Σ be an axiomatic system and let $\Delta \in \mathscr{W}_\Sigma$.

Suppose $\Psi_\Delta = L_V$. We will show that Ψ_Δ satisfies the definition of a cut around, Δ definition 5.1.1.

1. Since Ψ_Δ contains all sentences then Ψ_Δ contains $\bot$.

2. Note that there are no sentences B such that $B \notin \Psi_\Delta$. Therefore, for all sentences A and B such that $A \in \Psi_\Delta$ and $B \notin \Psi_\Delta$, it is vacuously the case that $(A \preccurlyeq B) \notin \Delta$.

Therefore Ψ_Δ is a cut around Δ. $\square$

We introduce some useful notation:

Definition 5.1.4.
Given a set of sentences Ψ, define $\neg\Psi$ as follows:

$$\neg\Psi := \{ \neg A \mid A \in \Psi \}.$$

This next theorem lists several properties about cuts.

Theorem 5.1.5.
Let Σ be an axiomatic system. Given a cut Ψ_Δ around a $\Delta \in \mathscr{W}_\Sigma$, then for any sentences A and B, the following hold:

1. $\Psi_\Delta = L_V$ *iff* $\neg\Psi_\Delta$ *is* Σ*-inconsistent.*
2. *If* $(A \prec B) \notin \Delta$ *and* $B \in \Psi_\Delta$ *then* $A \in \Psi_\Delta$.
3. *If* $\vdash_\Sigma A \to B$ *and* $B \in \Psi_\Delta$*, then* $A \in \Psi_\Delta$.
4. *If* $\vdash_\Sigma A$ *and* $A \in \Psi_\Delta$*, then* $\Psi_\Delta = L_V$.
5. *If* $A \notin \Psi_\Delta$ *then* $\{ A \} \cup \neg\Psi_\Delta$ *is* Σ*-consistent.*

Proof. Let Σ be an axiomatic system and let Ψ_Δ be cut around a $\Delta \in \mathscr{W}_\Sigma$. Let A and B be arbitrary sentences. We now prove each statement.

1. We prove each direction.

   For the first direction, we will show that

   $$\text{if } \Psi_\Delta = L_V \text{, then } \neg\Psi_\Delta \text{ is } \Sigma\text{-inconsistent.}$$

   Suppose $\Psi_\Delta = L_V$.

   Since Ψ_Δ contains all sentences, then there is some sentence A such that $\{ A, \neg A \} \subseteq \Psi_\Delta$. This implies that $\{ \neg A, \neg\neg A \} \subseteq \neg\Psi_\Delta$ and since $\{ \neg A, \neg\neg A \} \vdash^\ell_\Sigma \bot$ then it follows that $\neg\Psi_\Delta$ must be Σ-inconsistent.

   For the second direction, we will show that

   $$\text{if } \neg\Psi_\Delta \text{ is } \Sigma\text{-inconsistent, then } \Psi_\Delta = L_V.$$

Suppose that $\neg\Psi_\Delta$ is Σ-inconsistent. We will show that for all sentences, C, $C \in \Psi_\Delta$. Suppose for the sake of contradiction that there exists some C such that $C \notin \Psi_\Delta$.

Since $\neg\Psi_\Delta$ is Σ-inconsistent then so is $\{ C \} \cup \neg\Psi_\Delta$, by corollary 3.6.3. This tells us that for some finite set of sentences, $\{ A_1, A_2, \dots A_n \} \subseteq \Psi_\Delta$, we have the following derivation:

1.	$\vdash_\Sigma (\neg A_1 \wedge \neg A_2 \wedge \cdots \wedge \neg A_n \wedge C) \to \bot$	By 3.6.1
2.	$\vdash_\Sigma \neg(\neg A_1 \wedge \neg A_2 \wedge \cdots \wedge \neg A_n \wedge C)$	Defn of $\neg$
3.	$\vdash_\Sigma (A_1 \vee A_2 \vee \cdots \vee A_n \vee \neg C)$	Pl,2
4.	$\vdash_\Sigma C \to (A_1 \vee A_2 \vee \cdots \vee A_n)$	Pl,3
5.	$\vdash_\Sigma (A_1 \preccurlyeq C) \vee (A_2 \preccurlyeq C) \vee \cdots \vee (A_n \preccurlyeq C)$	Cp,4

Since this is a theorem in Σ then, by theorem 3.7.4(item 1), it must be in all maximally Σ-consistent sets, including Δ. That is to say,

$$((A_1 \preccurlyeq C) \vee (A_2 \preccurlyeq C) \vee \cdots \vee (A_n \preccurlyeq C)) \in \Delta. \tag{5.1}$$

Recall, that by assumption we have that Ψ_Δ is a cut. Moreover, since we assumed $C \notin \Psi_\Delta$ and we have $\{ A_1, A_2, \dots A_n \} \subseteq \Psi_\Delta$, then from the definition of a cut(5.1.1) it follows that for each $A_i \in \{ A_1, A_2, \dots A_n \}$ we have $\neg(A_i \preccurlyeq C) \in \Delta$. Since Δ is maximally Σ-consistent this gives us

$$(\neg(A_1 \preccurlyeq C) \wedge \neg(A_2 \preccurlyeq C) \wedge \cdots \wedge \neg(A_n \preccurlyeq C)) \in \Delta.$$

Taking this further we have

$$\neg((A_1 \preccurlyeq C) \vee (A_2 \preccurlyeq C) \vee \cdots \vee (A_n \preccurlyeq C)) \in \Delta. \tag{5.2}$$

This result (eq. (5.2)) contradicts our earlier finding (eq. (5.1)) given the fact that Δ is maximally Σ-consistent. Therefore, there does not exist a sentence C such that $C \notin \Psi_\Delta$, and hence $\Psi_\Delta = L_V$.

2. Suppose $(A \prec B) \notin \Delta$ and $B \in \Psi_\Delta$. For the sake of contradiction, suppose $A \notin \Psi_\Delta$.

   By the definition of a cut, definition 5.1.1, we obtain $(A \prec B) \in \Delta$. This is a contradiction and therefore $A \in \Psi_\Delta$.

3. Suppose $\vdash_\Sigma A \rightarrow B$ and $B \in \Psi_\Delta$. For the sake of contradiction, suppose $A \notin \Psi_\Delta$.

   Our assumption gives us the derivation:

   1. $\vdash_\Sigma A \rightarrow B$ By assumption
   2. $\vdash_\Sigma B \preccurlyeq A$ CP, 1

   Since this is a theorem of Σ then it is in all maximally Σ-consistent sets by theorem 3.7.4(item 1), and therefore $(B \preccurlyeq A) \in \Delta$. Then, since Δ is maximally Σ-consistent that means that $(A \prec B) \notin \Delta$, and by theorem 5.1.5(item 2) we conclude that $A \in \Psi_\Delta$.

4. Suppose $\vdash_\Sigma A$ and $A \in \Psi_\Delta$.

   Then, for every sentence B we have the theorem $\vdash_\Sigma B \rightarrow A$, by PL, and by theorem 5.1.5(item 3) we have that $B \in \Psi_\Delta$. Therefore Ψ_Δ contains all sentences.

5. Suppose $A \notin \Psi_\Delta$. Then, $\Psi_\Delta \neq L_V$ and by theorem 5.1.5(item 1) it follows that $\neg\Psi_\Delta$ is Σ-consistent.

   Next, Suppose, for the sake of contradiction that $\{ A \} \cup \neg\Psi_\Delta$ is not Σ-consistent,

   $$\{ A \} \cup \neg\Psi_\Delta \vdash^{\ell}_{\Sigma} \bot.$$

   Meaning that for some sentences $\{ A_1, A_2, \dots A_n \} \subseteq \Psi_\Delta$ we obtain the following derivation:

   1. $\vdash_\Sigma (\neg A_1 \wedge \neg A_2 \wedge \cdots \wedge \neg A_n \wedge A) \rightarrow \bot$ By definition 3.6.1
   2. $\vdash_\Sigma \neg(\neg A_1 \wedge \neg A_2 \wedge \cdots \wedge \neg A_n \wedge A)$ Defn of $\neg$
   3. $\vdash_\Sigma (A_1 \vee A_2 \vee \cdots \vee A_n \vee \neg A)$ PL, 2
   4. $\vdash_\Sigma A \rightarrow (A_1 \vee A_2 \vee \cdots \vee A_n)$ PL, 3
   5. $\vdash_\Sigma (A_1 \preccurlyeq A) \vee (A_2 \preccurlyeq A) \vee \cdots \vee (A_n \preccurlyeq A)$ CP, 4

   Since this is a theorem in Σ then, by theorem 3.7.4(item 1), it must be true in all maximally Σ-consistent sets, including Δ. That is to say,

   $$((A_1 \preccurlyeq A) \vee (A_2 \preccurlyeq A) \vee \cdots \vee (A_n \preccurlyeq A)) \in \Delta. \tag{5.3}$$

Moreover, since we assumed $A \notin \Psi_\Delta$ and we have $\{ A_1, A_2, \dots A_n \} \subseteq \Psi_\Delta$, then from the definition of a cut, definition 5.1.1, it follows that for each $A_i \in \{ A_1, A_2, \dots A_n \}$ we have $\neg(A_i \preccurlyeq A) \in \Delta$. Since Δ is maximally Σ-consistent this gives us

$$(\neg(A_1 \preccurlyeq A) \wedge \neg(A_2 \preccurlyeq A) \wedge \cdots \wedge \neg(A_n \preccurlyeq A)) \in \Delta.$$

Taking this further we have

$$\neg((A_1 \preccurlyeq A) \vee (A_2 \preccurlyeq A) \vee \cdots \vee (A_n \preccurlyeq A)) \in \Delta. \tag{5.4}$$

This result (eq. (5.3)) together with our earlier finding (eq. (5.4)) directly contradicts the fact that Δ is maximally Σ-consistent. Therefore, our assumption must be false and we conclude that if $A \notin \Psi_\Delta$ then $\{ A \} \cup \neg\Psi_\Delta$ is Σ-consistent.

□

Now we can define a co-sphere. Note that for each cut we only obtain one co-sphere, this allows us to formulate this notion with a function.

Definition 5.1.6 (Co-sphere of Ψ_Δ)**.**
Given an axiomatic system, Σ, and a cut Ψ_Δ around a $\Delta \in \mathscr{W}_\Sigma$ we define the **co-sphere of** Ψ_Δ to be the set of all and only those $\Gamma \in \mathscr{W}_\Sigma$ that contain $\neg\Psi_\Delta$ as a subset.

We may further define a function that assigns to each cut around Δ its corresponding co-sphere around Δ:

$$\begin{aligned} \mathscr{S} &: \textsc{Cuts}(\Delta) \to \mathscr{P}(\mathscr{W}_\Sigma) \\ \mathscr{S} &: \Psi_\Delta \mapsto \left\{ \Gamma \in \mathscr{W}_\Sigma \;\middle|\; \neg\Psi_\Delta \subseteq \Gamma \right\}. \end{aligned}$$

Remark 5.1.7. In *Counterfactuals* (Lewis, 2001, Section 6.1 p. 127), Lewis defines a co-sphere of a cut Ψ_Δ around a $\Delta \in \mathscr{W}_\Sigma$, to be the set of all and only those $\Gamma \in \mathscr{W}_\Sigma$ that contain no sentence in Ψ_Δ. In other words, given a cut Ψ_Δ, the corresponding co-sphere is the set

$$\left\{ \Gamma \in \mathscr{W}_\Sigma \;\middle|\; \forall A \in \Psi_\Delta, A \notin \Gamma \right\}.$$

The equivalence between our definition of a co-sphere, definition 5.1.6, and Lewis' definition of a co-sphere remark 5.1.7 (Lewis, 2001, Section 6.1 p. 127) follows from the following lemma.

Lemma 5.1.8.
Given an axiomatic system, Σ, and a cut Ψ_Δ around a $\Delta \in \mathscr{W}_\Sigma$. For all $\Gamma \in \mathscr{W}_\Sigma$, the following are equivalent

1. $\neg\Psi_\Delta \subseteq \Gamma$.

2. $\forall A \in \Psi_\Delta, A \notin \Gamma$.

Proof. Let Σ be an axiomatic system and let Ψ_Δ be a cut around a $\Delta \in \mathscr{W}_\Sigma$. Let $\Gamma \in \mathscr{W}_\Sigma$ be arbitrary.

We'll do the proof in two directions.

First we prove that $\neg\Psi_\Delta \subseteq \Gamma$ implies $\forall A \in \Psi_\Delta, A \notin \Gamma$. Suppose $\neg\Psi_\Delta \subseteq \Gamma$. Let $A \in \Psi_\Delta$ be arbitrary, then $\neg A \in \neg\Psi_\Delta$ and therefore $\neg A \in \Gamma$. Then since Γ is maximally Σ-consistent, it follows that $A \notin \Gamma$, and since A was arbitrary we get $\forall A \in \Psi_\Delta, A \notin \Gamma$, which completes the first direction of our proof.

Next we prove that $\forall A \in \Psi_\Delta, A \notin \Gamma$ implies $\neg\Psi_\Delta \subseteq \Gamma$. Suppose $\forall A \in \Psi_\Delta, A \notin \Gamma$. Let $\neg A \in \neg\Psi_\Delta$, then by definition of $\neg\Psi_\Delta$ it follows that $A \in \Psi_\Delta$. Then, by our assumption, we get that $A \notin \Gamma$, and since Γ is maximally Σ-consistent this tells us that $\neg A \in \Gamma$. Since $\neg A \in \neg\Psi_\Delta$ was arbitrary, then we get $\neg\Psi_\Delta \subseteq \Gamma$, which completes the second direction of our proof. □

5.2 Canonical Models

Finally, we can introduce the definition of the canonical model.

Definition 5.2.1 (Canonical Model (Lewis, 2001, Section 6.1))**.**
Given an axiomatic system, Σ. The canonical model for Σ is $M_\Sigma = (\mathscr{W}_\Sigma, \mathcal{O}, v)$, where:

1. $\mathscr{W}_\Sigma$ is the set of all maximally Σ-consistent sets of sentences, as previously defined, definition 3.7.3.

2. $\mathscr{O}$ is a system of spheres, called the *canonical basis*, such that for each $\Delta \in \mathscr{W}_\Sigma$, $\mathscr{O}(\Delta)$ consists of all the unions of sets of co-spheres of cuts around Δ.[1]

$$\mathscr{O} : \mathscr{W}_\Sigma \to \mathscr{P}(\mathscr{P}(\mathscr{W}_\Sigma))$$
$$\mathscr{O} : \Delta \mapsto \left\{ \bigcup_{\Psi_\Delta \in X} \mathcal{S}(\Psi_\Delta) \;\middle|\; X \subseteq \textsc{Cuts}(\Delta) \right\}.$$

3. v is the valuation function:

$$v : \textsc{Props} \to \mathscr{P}(\mathscr{W}_\Sigma)$$
$$v : p \mapsto \left\{ \Delta \in \mathscr{W}_\Sigma \;\middle|\; p \in \Delta \right\}.$$

We now need to show that the canonical model satisfies the definition of a sphere model. To do this we only need to prove that the canonical basis is indeed a system of spheres. That is to say, we must prove that for each $\Delta \in \mathscr{W}_\Sigma$, $\mathscr{O}(\Delta)$ is totally ordered by $\subseteq$, and $\mathscr{O}(\Delta)$ is closed under arbitrary union and non-empty intersection.

Our plan for this is as follows:

1. Prove that the set of cuts around a given Δ, $\textsc{Cuts}(\Delta)$, is:
   a) Totally ordered by $\subseteq$.
   b) Closed under non-empty union.
   c) Closed under arbitrary intersection.
2. Prove that the set of co-spheres of cuts around a given Δ, $\mathcal{S}(\textsc{Cuts}(\Delta))$, is:
   a) Totally ordered by $\subseteq$ (follows from 1a).
   b) Closed under non-empty intersection (follows from 1b).
   c) **Not** closed under union, but union yields a subset (follows from 1c). We don't need this result to show that it is a sphere model, but we prove it because it is closely related to an important counter-example we'll discuss later.

[1] We show later that unions of sets of co-spheres (of cuts around Δ) are not necessarily co-spheres (of cuts around Δ) themselves.

3. Prove that the set of spheres, $\mathscr{O}(\Delta)$, is:
   a) Totally ordered by $\subseteq$ (follows from 2a).
   b) Closed under arbitrary union (trivial).
   c) Closed under non-empty intersection (follows from 2b).

We'll be dealing with a lot of unions and intersections throughout this section, so it will be helpful to prove this lemma that lets us move negations in and out of unions and intersections of cuts.

Lemma 5.2.2.
Given an axiomatic system, Σ, and a maximally Σ-consistent set Δ, if X is a non-empty set of cuts around Δ, we have the following:

1.
$$\neg \bigcup_{\Psi_\Delta \in X} \Psi_\Delta = \bigcup_{\Psi_\Delta \in X} \neg \Psi_\Delta.$$

2.
$$\neg \bigcap_{\Psi_\Delta \in X} \Psi_\Delta = \bigcap_{\Psi_\Delta \in X} \neg \Psi_\Delta.$$

Proof. Let Σ be an axiomatic system and Δ be a maximally Σ-consistent set. Suppose X is an arbitrary non-empty set of cuts around Δ.

Recall that for all sets of sentences, Ψ,

$$\neg A \in \neg \Psi \text{ iff } A \in \Psi.$$

This tells us that $\neg$ is injective when interpreted as a function and that every sentence in a negated set is itself negated. We use both remarks to streamline the arguments below but injectivity is only required for the second claim.

We prove each statement separately:

1. The proof is as follows.

$$\begin{aligned}\neg A \in \neg \bigcup_{\Psi_\Delta \in X} \Psi_\Delta \\ &\text{iff } A \in \bigcup_{\Psi_\Delta \in X} \Psi_\Delta \\ &\text{iff } \exists \Psi_\Delta \in X, A \in \Psi_\Delta \\ &\text{iff } \exists \Psi_\Delta \in X, \neg A \in \neg \Psi_\Delta \\ &\text{iff } \neg A \in \bigcup_{\Psi_\Delta \in X} \neg \Psi_\Delta .\end{aligned}$$

2. This lemma requires the use of injectivity in the reverse direction but otherwise it is the same as above. We do not explicitly show the use of injectivity for neatness.

$$\begin{aligned}\neg A \in \neg \bigcap_{\Psi_\Delta \in X} \Psi_\Delta \\ &\text{iff } A \in \bigcap_{\Psi_\Delta \in X} \Psi_\Delta \\ &\text{iff } \forall \Psi_\Delta \in X, A \in \Psi_\Delta \\ &\text{iff } \forall \Psi_\Delta \in X, \neg A \in \neg \Psi_\Delta \\ &\text{iff } \neg A \in \bigcap_{\Psi_\Delta \in X} \neg \Psi_\Delta .\end{aligned}$$

□

We proceed now with the first step of our plan. Showing that the cuts around a maximally Σ-consistent set are totally ordered, closed under non-empty union, and closed under arbitrary intersection.

Theorem 5.2.3.
Given an axiomatic system, Σ, and a maximally Σ-consistent $\Delta \in \mathscr{W}_\Sigma$. the following are true:

1. *Cuts(Δ) is totally ordered by ⊆:*

   for every $\Psi_\Delta, \Phi_\Delta \in$ Cuts(Δ), either $\Psi_\Delta \subseteq \Phi_\Delta$ or $\Phi_\Delta \subseteq \Psi_\Delta$.

2. *Cuts*(Δ) *is closed under non-empty union:*

   if X is a non-empty set of cuts around Δ, then $\bigcup_{\Psi_\Delta \in X} \Psi_\Delta$ is a cut around Δ.

3. *Cuts*(Δ) *is closed under arbitrary intersection:*

   if X is a set of cuts around Δ, then $\bigcap_{\Psi_\Delta \in X} \Psi_\Delta$ is a cut around Δ.

Proof. Let Σ be an axiomatic system and let $\Delta \in \mathscr{W}_\Sigma$ be a maximally Σ-consistent set.

We prove each statement separately:

1. Suppose for the sake of contradiction that there exist two cuts, Ψ_Δ and Φ_Δ around Δ, such that

   $$\Psi_\Delta \nsubseteq \Phi_\Delta \text{ and } \Phi_\Delta \nsubseteq \Psi_\Delta.$$

   Then there exists a sentence $B \in \Psi_\Delta$ such that $B \notin \Phi_\Delta$ and a sentence $A \in \Phi_\Delta$ such that $A \notin \Psi_\Delta$.

   Next, since they are both cuts around Δ, it follows from the definition of a cut, remark 5.1.2, that $\neg(B \preccurlyeq A) \in \Delta$ and $\neg(A \preccurlyeq B) \in \Delta$. However, since

   $$\vdash_\Sigma (\neg(B \preccurlyeq A) \wedge \neg(A \preccurlyeq B)) \rightarrow \neg((B \preccurlyeq A) \vee (A \preccurlyeq B)),$$

   then, by definition 3.5.1, $\Delta \vdash^{\ell}_{\Sigma} \neg((B \preccurlyeq A) \vee (A \preccurlyeq B))$. Then, since Δ is maximally Σ-consistent, by deductive closure theorem 3.7.4(item 1) we get that $\neg((B \preccurlyeq A) \vee (A \preccurlyeq B)) \in \Delta$. However, since strong connectivity tells us that $\vdash_\Sigma ((B \preccurlyeq A) \vee (A \preccurlyeq B))$ and therefore in $((B \preccurlyeq A) \vee (A \preccurlyeq B)) \in \Delta$ then this contradicts the fact that Δ is maximally Σ-consistent. Thus, by contradiction, no such cuts can exist.

   Therefore it must be the case that Cuts(Δ) is totally ordered by $\subseteq$.

2. Suppose X is an arbitrary non-empty set of cuts around Δ. We want to show that $\bigcup_{\Psi_\Delta \in X} \Psi_\Delta$ is itself a cut around Δ by showing that it satisfies the definition of a cut around Δ.

   a) Since $\perp$ is contained in all cuts around Δ, then for all $\Psi_\Delta \in X$, we have $\perp \in \Psi_\Delta$. Therefore, $\perp \in \bigcup_{\Psi_\Delta \in X} \Psi_\Delta$.

b) Let A and B be arbitrary sentences. Suppose $B \in \bigcup_{\Psi_\Delta \in X} \Psi_\Delta$ and $A \notin \bigcup_{\Psi_\Delta \in X} \Psi_\Delta$. Then there must exist a cut around Δ, $\Psi_\Delta \in X$, such that $B \in \Psi_\Delta$ but $A \notin \Psi_\Delta$. Let Ψ_Δ be such a cut, then by the definition of a cut around Δ, definition 5.1.1, we have $(A \prec B) \in \Delta$. Thus, since A and B were arbitrary, we have that for all sentences, A and B, if $B \in \bigcup_{\Psi_\Delta \in X} \Psi_\Delta$ and $A \notin \bigcup_{\Psi_\Delta \in X} \Psi_\Delta$ then $(A \prec B) \in \Delta$.

Therefore, $\bigcup_{\Psi_\Delta \in X} \Psi_\Delta$ satisfies the definition of a cut around Δ, definition 5.1.1, and hence $\textsc{Cuts}(\Delta)$ is closed under non-empty union.

3. Suppose X is an arbitrary set of cuts around Δ. In the case where $X = \varnothing$ then $\bigcup_{\Psi_\Delta \in X} \Psi_\Delta = L_V$ which is a cut by theorem 5.1.3. So, consider the case where X is non-empty. We want to show that $\bigcap_{\Psi_\Delta \in X} \Psi_\Delta$ is itself a cut around Δ by showing that it satisfies the definition of a cut around Δ.

   a) Since $\bot$ is contained in all cuts then for all $\Psi_\Delta \in X$ we have $\bot \in \Psi_\Delta$. Therefore, $\bot \in \bigcap_{\Psi_\Delta \in X} \Psi_\Delta$.

   b) Let A and B be arbitrary sentences. Suppose that $B \in \bigcap_{\Psi_\Delta \in X} \Psi_\Delta$ but $A \notin \bigcap_{\Psi_\Delta \in X} \Psi_\Delta$. Then (since we're assuming X is non-empty) there must exist a cut around Δ, $\Psi_\Delta \in X$, such that $B \in \Psi_\Delta$ but $A \notin \Psi_\Delta$. Let Ψ_Δ be such a cut, then by the definition of a cut around Δ, definition 5.1.1, this implies that $(A \prec B) \in \Delta$. Thus, we've shown that, for all sentences A and B, if $B \in \bigcap_{\Psi_\Delta \in X} \Psi_\Delta$ but $A \notin \bigcap_{\Psi_\Delta \in X} \Psi_\Delta$ then, $(A \prec B) \in \Delta$.

   Therefore, $\bigcap_{\Psi_\Delta \in X} \Psi_\Delta$ satisfies the definition of a cut around Δ, definition 5.1.1, and hence $\textsc{Cuts}(\Delta)$ is closed under arbitrary intersection.

□

Next, we continue on with the second step of our plan, showing that co-spheres of cuts around a maximally Σ-consistent set are totally ordered and closed under non-empty intersection. We do not show that they are closed under arbitrary union, but we do show that taking the arbitrary union of co-spheres gives us a subset of of a co-sphere. If this last one seems odd, know that it is. We'll discuss it more later.

Theorem 5.2.4.
Given an axiomatic system, Σ, and a maximally Σ-consistent $\Delta \in \mathscr{W}_\Sigma$, the following are true:

1. *$\mathcal{S}(\textsc{Cuts}(\Delta))$ is totally ordered by $\subseteq$:*

   for every $\mathcal{S}(\Psi_\Delta), \mathcal{S}(\Phi_\Delta) \in \mathcal{S}(\textsc{Cuts}(\Delta))$, either $\mathcal{S}(\Psi_\Delta) \subseteq \mathcal{S}(\Phi_\Delta)$ or $\mathcal{S}(\Phi_\Delta) \subseteq \mathcal{S}(\Psi_\Delta)$.

2. *$\mathcal{S}(\textsc{Cuts}(\Delta))$ is closed under non-empty intersection:*

   if X is a non-empty set of cuts around Δ,

$$\bigcap_{\Psi_\Delta \in X} \mathcal{S}\left(\Psi_\Delta\right) = \mathcal{S}\left(\bigcup_{\Psi_\Delta \in X} \Psi_\Delta\right).$$

3. *The arbitrary union of a set of co-spheres in $\mathcal{S}(\textsc{Cuts}(\Delta))$ is a subset of a co-sphere in $\mathcal{S}(\textsc{Cuts}(\Delta))$:*

   if X is a non-empty set of cuts around Δ,

$$\bigcup_{\Psi_\Delta \in X} \mathcal{S}\left(\Psi_\Delta\right) \subseteq \mathcal{S}\left(\bigcap_{\Psi_\Delta \in X} \Psi_\Delta\right).$$

Proof. Let Σ be an axiomatic system and let $\Delta \in \mathscr{W}_\Sigma$ be a maximally Σ-consistent set.

We prove each statement separately:

1. Let $\mathcal{S}(\Psi_\Delta), \mathcal{S}(\Phi_\Delta)$ be arbitrary co-spheres in $\mathcal{S}(\textsc{Cuts}(\Delta))$.

   By theorem 5.2.3(item 1) we know that $\textsc{Cuts}(\Delta)$ is totally ordered by $\subseteq$, so, suppose without loss of generality that $\Psi_\Delta \subseteq \Phi_\Delta$. Then it follows that $\neg\Psi_\Delta \subseteq \neg\Phi_\Delta$.

Next, recall the definition of co-sphere, definition 5.1.6,

$$\mathcal{S}(\Psi_\Delta) = \left\{ \Gamma \in \mathscr{W}_\Sigma \;\middle|\; \neg\Psi_\Delta \subseteq \Gamma \right\},$$

and let $\Gamma \in \mathcal{S}(\Phi_\Delta)$ be arbitrary, i.e., $\Gamma \in \mathscr{W}_\Sigma$ and $\neg\Phi_\Delta \subseteq \Gamma$. However, since $\neg\Psi_\Delta \subseteq \neg\Phi_\Delta$ then $\neg\Psi_\Delta \subseteq \Gamma$ and therefore $\Gamma \in \mathcal{S}(\Psi_\Delta)$. Since Γ was arbitrary, then $\mathcal{S}(\Phi_\Delta) \subseteq \mathcal{S}(\Psi_\Delta)$.

Therefore, it follows that the set of co-spheres around Δ is totally ordered by $\subseteq$.

2. Suppose X is an arbitrary non-empty set of cuts around Δ.

$$\Gamma \in \bigcap_{\Psi_\Delta \in X} \mathcal{S}\left(\Psi_\Delta\right)$$

$$\text{iff } \forall\Psi_\Delta \in X, \Gamma \in \mathcal{S}(\Psi_\Delta)$$

$$\text{iff } \forall\Psi_\Delta \in X, (\Gamma \in \mathscr{W}_\Sigma \text{ and } \neg\Psi_\Delta \subseteq \Gamma)$$

(by definition 5.1.6)

$$\text{iff } \Gamma \in \mathscr{W}_\Sigma \text{ and } \forall\Psi_\Delta \in X, \neg\Psi_\Delta \subseteq \Gamma$$

$$\text{iff } \Gamma \in \mathscr{W}_\Sigma \text{ and } \left(\bigcup_{\Psi_\Delta \in X} \neg\Psi_\Delta\right) \subseteq \Gamma$$

$$\text{iff } \Gamma \in \mathscr{W}_\Sigma \text{ and } \neg\left(\bigcup_{\Psi_\Delta \in X} \Psi_\Delta\right) \subseteq \Gamma$$

(by lemma 5.2.2)

$$\text{iff } \Gamma \in \mathcal{S}\left(\bigcup_{\Psi_\Delta \in X} \Psi_\Delta\right)$$

(by definition 5.1.6 and theorem 5.2.3(item 2)).

Therefore,

$$\bigcap_{\Psi_\Delta \in X} \mathcal{S}\left(\Psi_\Delta\right) = \mathcal{S}\left(\bigcup_{\Psi_\Delta \in X} \Psi_\Delta\right),$$

and hence, $\mathcal{S}(\text{CUTS}(\Delta))$ is closed under non-empty intersection.

3. Suppose X is an arbitrary non-empty set of cuts around Δ.

$$\Gamma \in \bigcup_{\Psi_\Delta \in X} \mathscr{S}\left(\Psi_\Delta\right)$$

$$\text{iff } \exists \Psi_\Delta \in X, \Gamma \in \mathscr{S}(\Psi_\Delta)$$

$$\text{iff } \exists \Psi_\Delta \in X, \left(\Gamma \in \mathscr{W}_\Sigma \text{ and } \neg\Psi_\Delta \subseteq \Gamma\right)$$

(by definition 5.1.6)

$$\text{iff } \Gamma \in \mathscr{W}_\Sigma \text{ and } \exists \Psi_\Delta \in X, \neg\Psi_\Delta \subseteq \Gamma \tag{5.5}$$

$$\textbf{only if } \Gamma \in \mathscr{W}_\Sigma \text{ and } \left(\bigcap_{\Psi_\Delta \in X} \neg\Psi_\Delta\right) \subseteq \Gamma \tag{5.6}$$

$$\text{iff } \Gamma \in \mathscr{W}_\Sigma \text{ and } \neg\left(\bigcap_{\Psi_\Delta \in X} \Psi_\Delta\right) \subseteq \Gamma$$

(by lemma 5.2.2)

$$\text{iff } \Gamma \in \mathscr{S}\left(\bigcap_{\Psi_\Delta \in X} \Psi_\Delta\right)$$

(by definition 5.1.6 and theorem 5.2.3(item 3)).

□

Remark 5.2.5. In the proof of theorem 5.2.4(item 3) it is crucially important that the "only if" step between eqs. (5.5) and (5.6) only works in one direction. If this proof held in both directions then it would imply that spheres are in fact co-spheres. Initially, Lewis had actually defined co-spheres as spheres in his first edition of *Counterfactuals* (Lewis, 1973) but Krabbe (1978) showed that this was flawed. Krabbe did this by constructing a counterexample where there exists a Γ such that $\Gamma \in \mathscr{S}\left(\bigcap_{\Psi_\Delta \in X} \Psi_\Delta\right)$ but $\Gamma \notin \bigcup_{\Psi_\Delta \in X} \mathscr{S}\left(\Psi_\Delta\right)$ (see theorem 7.4.3). Moreover, he showed that $\bigcup_{\Psi_\Delta \in X} \mathscr{S}\left(\Psi_\Delta\right)$ is not the co-sphere of any cut around Δ (see theorem 7.4.5). We dedicate an entire chapter later on to presenting the details of Krabbe's argument.

Finally we reach the last step in our plan. Proving that the spheres around a maximally Σ-consistent set are totally ordered, closed under arbitrary union, and closed under non-empty intersection.

Theorem 5.2.6.
Given an axiomatic system, Σ, and a maximally Σ-consistent $\Delta \in \mathscr{W}_\Sigma$, the following hold:

1. *$\mathcal{O}(\Delta)$ is totally ordered by $\subseteq$:*

   for every, $S_1, S_2 \in \mathcal{O}(\Delta)$, either $S_1 \subseteq S_2$ or $S_2 \subseteq S_1$.

2. *$\mathcal{O}(\Delta)$ is closed under arbitrary union:*

   If $O \subseteq \mathcal{O}(\Delta)$ is an arbitrary set of spheres around Δ, then

$$\left(\bigcup_{S \in O} S \right) \in \mathcal{O}(\Delta).$$

3. *$\mathcal{O}(\Delta)$ is closed under non-empty intersection:*

   if $O \subseteq \mathcal{O}(\Delta)$ is a non-empty set of spheres around Δ, then

$$\left(\bigcap_{S \in O} S \right) \in \mathcal{O}(\Delta).$$

Proof. Let Σ be an axiomatic system and let $\Delta \in \mathscr{W}_\Sigma$ be a maximally Σ-consistent set.

1. Let $S_1, S_2 \in \mathcal{O}(\Delta)$ be arbitrary. If $S_1 = S_2$ then we are done, so suppose they are distinct. Then, since they are distinct, there must exist an element in one that is not in the other. Without loss of generality, suppose $\Gamma \in S_1$ but $\Gamma \notin S_2$.

   Since S_1 and S_2 are, by definition, unions of co-spheres (of cuts around Δ), then there exist sets of cuts around Δ, X_1 and X_2, such that

$$S_1 = \bigcup_{\Psi_\Delta \in X_1} \mathcal{S}(\Psi_\Delta) \qquad S_2 = \bigcup_{\Phi_\Delta \in X_2} \mathcal{S}(\Phi_\Delta).$$

First, observe that:

$$\begin{aligned}
&\Gamma \in S_1 \\
&\qquad \text{iff } \Gamma \in \bigcup_{\Psi_\Delta \in X_1} \mathcal{S}(\Psi_\Delta) \\
&\qquad \text{iff } \exists \Psi_\Delta \in X_1, \Gamma \in \mathcal{S}(\Psi_\Delta).
\end{aligned}$$

Next, observe that:

$$\begin{aligned}
&\Gamma \notin S_2 \\
&\qquad \text{iff } \Gamma \notin \bigcup_{\Phi_\Delta \in X_2} \mathcal{S}(\Phi_\Delta) \\
&\qquad \text{iff } \forall \Phi_\Delta \in X_2, \Gamma \notin \mathcal{S}(\Phi_\Delta).
\end{aligned}$$

Let $\Psi_\Delta \in X_1$ be a cut such that $\Gamma \in \mathcal{S}(\Psi_\Delta)$ and let $\Phi_\Delta \in X_2$ be arbitrary. Then, because all co-sphere are nested by theorem 5.2.4(item 1), it must be the case that either

$$\mathcal{S}(\Psi_\Delta) \subseteq \mathcal{S}(\Phi_\Delta) \quad \text{or} \quad \mathcal{S}(\Phi_\Delta) \subseteq \mathcal{S}(\Psi_\Delta).$$

If we suppose $\mathcal{S}(\Psi_\Delta) \subseteq \mathcal{S}(\Phi_\Delta)$, then since $\Gamma \in \mathcal{S}(\Psi_\Delta)$ it follows that $\Gamma \in \mathcal{S}(\Phi_\Delta)$ which is a contradiction. Therefore, it must be the case that $\mathcal{S}(\Phi_\Delta) \subseteq \mathcal{S}(\Psi_\Delta)$. Moreover, since $\Phi_\Delta \in X_2$ was arbitrary, then $\bigcup_{\Phi_\Delta \in X_2} \mathcal{S}(\Phi_\Delta) \subseteq \mathcal{S}(\Psi_\Delta)$. Finally, by noting the definition of S_2 and the fact that $\mathcal{S}(\Psi_\Delta) \subseteq S_1$, we get

$$S_2 \subseteq S_1.$$

Therefore, we have that $\mathcal{O}(\Delta)$ must be totally ordered by $\subseteq$.

2. The result follows trivially from the definition of $\mathcal{O}(\Delta)$.

3. Suppose $O \subseteq \mathcal{O}(\Delta)$ is a non-empty set of spheres around Δ.

   If $\varnothing \in O$ then we're done, since the intersection is the empty set. So, suppose $\varnothing \notin O$.

Recall that by the definition of a sphere, each $S \in O$ is a union of co-spheres, meaning that for each S there exists a set of cuts X_S such that

$$S = \bigcup_{\Psi_\Delta \in X_S} \mathcal{S}(\Psi_\Delta).$$

where each $\mathcal{S}(\Psi_\Delta) \in \mathcal{O}(\Delta)$ by definition.

First we'll establish the existence of choice functions. Let $S \in O$ be arbitrary. Suppose $X_S = \varnothing$. That would mean that $S = \varnothing$ because it's an empty union. This contradicts our assumption and therefore, $\forall S \in O$, $X_S \neq \varnothing$. Then, note by the axiom of choice (Dasgupta, 2014, Section 5.5, p. 94), since $\{ X_S \mid S \in O \}$ is an indexed family of sets with $X_S \neq \varnothing$ for all $S \in O$ then there exists a choice function $f : O \to \bigcup_{S \in O} X_S$ such that $f(S) \in X_S$ for all $S \in O$.

Let

$$F = \left\{ f \;\middle|\; f : O \to \bigcup_{S \in O} X_S \text{ such that } \forall S \in O, f(S) \in X_S \right\},$$

be the set of such choice functions.

So,

$$\begin{aligned}
\bigcap_{S \in O} S &= \bigcap_{S \in O} \left(\bigcup_{\Psi_\Delta \in X_S} \mathcal{S}(\Psi_\Delta) \right) \\
&= \bigcup_{f \in F} \left(\bigcap_{S \in O} \mathcal{S}(f(S)) \right) \\
&= \bigcup_{f \in F} \left(\mathcal{S} \left(\bigcup_{S \in O} f(S) \right) \right)
\end{aligned}$$

(by theorem 5.2.4(item 2)).

Then, since for each $f \in F$, and each $S \in O$, $f(S)$ is a cut around Δ, then it follows from theorem 5.2.3(item 2) that $\bigcup_{S \in O} f(S)$ is a cut around Δ. Therefore, $\mathcal{S}\left(\bigcup_{S \in O} f(S)\right)$ is a co-sphere of a cut around Δ, and thus by definition,

$$\bigcup_{f \in F} \left(\mathcal{S} \left(\bigcup_{S \in O} f(S) \right) \right) \in \mathcal{O}(\Delta).$$

since, $\mathcal{O}(\Delta)$ is precisely the set of unions of sets of co-spheres. Hence,

$$\bigcap_{S \in O} S \in \mathcal{O}(\Delta),$$

as desired.

□

With these proofs done it's easy to conclude that canonical basis is in fact a system of spheres.

Corollary 5.2.7.
The canonical basis, $\mathcal{O}$, is a system of spheres.

Proof. Theorem 5.2.6 proves that $\mathcal{O}$ satisfies all of the properties of a system of spheres, definition 2.2.1. □

Therefore, our canonical model is indeed a sphere model.

5.3 Truth Lemma

We need to show that in the canonical model, for all sentences, C, and $\forall \Gamma \in \mathscr{W}_{\Sigma}$, $\Gamma \in [\![C]\!]$ iff $C \in \Gamma$. This is called the Truth Lemma. Lewis does this using what he calls the co-sphere lemma. Unfortunately, there is a very nuanced problem with his approach where his co-sphere lemma relies very subtly on the Truth lemma and the truth lemma itself must be proven by induction. To resolve this, we present a slightly weaker version of the co-sphere lemma, which we dub the "weak co-sphere lemma." Then, we use this lemma in the inductive step of the truth theorem, and finally we prove Lewis' version of the co-sphere lemma at the end, so that it is present.

Lemma 5.3.1 (Weak Co-Sphere Lemma)**.**
Given the canonical model, $M_\Sigma = (\mathscr{W}_\Sigma, \mathscr{O}, v)$ for an axiomatic system, Σ, and given a sentence A, if $\mathcal{S}(\Psi_\Delta)$ is the co-sphere of a cut around a $\Delta \in \mathscr{W}_\Sigma$, then $A \in \Psi_\Delta$ iff $\mathcal{S}(\Psi_\Delta) \not\bullet \left\{ \Theta \in \mathscr{W}_\Sigma \;\middle|\; A \in \Theta \right\}$.

Proof. Let $M_\Sigma = (\mathscr{W}_\Sigma, \mathscr{O}, v)$ be a canonical model for an axiomatic system, Σ. Let A be an arbitrary sentence. Suppose $\mathcal{S}(\Psi_\Delta)$ is the co-sphere of a cut around a $\Delta \in \mathscr{W}_\Sigma$.

We want to show that $A \in \Psi_\Delta$ iff $\mathcal{S}(\Psi_\Delta) \not\bullet \left\{ \Theta \in W_\Sigma \;\middle|\; A \in \Theta \right\}$.

Suppose $A \in \Psi_\Delta$. Then $\neg A \in \neg\Psi_\Delta$. Let $\Gamma \in \mathcal{S}(\Psi_\Delta)$ be arbitrary. Then $\neg A \in \Gamma$ and therefore $A \notin \Gamma$. Moreover, $\Gamma \notin \left\{ \Theta \in W_\Sigma \;\middle|\; A \in \Theta \right\}$, meaning that $\mathcal{S}(\Psi_\Delta) \not\bullet \left\{ \Theta \in W_\Sigma \;\middle|\; A \in \Theta \right\}$.

Suppose $A \notin \Psi_\Delta$. Then $\{ A \} \cup \neg\Psi_\Delta$ must be Σ-consistent. Extend this to a maximally Σ-consistent, Γ. Then $\Gamma \in \mathcal{S}(\Psi_\Delta)$ by definition. Moreover, $A \in \Gamma$ by construction meaning that $\mathcal{S}(\Psi_\Delta) \bullet \left\{ \Theta \in W_\Sigma \;\middle|\; A \in \Theta \right\}$. □

We will make use of the following result within the proof for the truth lemma.

Lemma 5.3.2.
Let Σ be an axiomatic system. Let $\Gamma \in \mathscr{W}_\Sigma$ be arbitrary.

The following set is a cut around Γ:

$$\Psi_\Gamma = \{ X \mid (A \preccurlyeq X) \in \Gamma \}.$$

Proof. Let Σ be an axiomatic system. Let $\Gamma \in \mathscr{W}_\Sigma$ be arbitrary. Let

$$\Psi_\Gamma = \{ X \mid (A \preccurlyeq X) \in \Gamma \}.$$

We now prove that Ψ_Γ satisfies the definition of a cut around Γ.

1. First, note that we have the following derivation:

   1. $\vdash_\Sigma \bot \to A$ Taut
   2. $\vdash_\Sigma (A \preccurlyeq \bot)$ Cp

   Then, by theorem 3.7.4(item 1) it follows that $(A \preccurlyeq \bot) \in \Gamma$. Therefore by the definition of Ψ_Γ we have that $\bot \in \Psi_\Gamma$.

2. Let C and D be arbitrary sentences such that $C \in \Psi_\Gamma$ and $D \notin \Psi_\Gamma$. From the definition of Ψ_Γ we get that $(A \preccurlyeq C) \in \Gamma$ but $(A \preccurlyeq D) \notin \Gamma$. Next, suppose for the sake of contradiction that $(C \preccurlyeq D) \in \Gamma$. Due to Trans, we have $\vdash_\Sigma ((A \preccurlyeq C) \wedge (C \preccurlyeq D)) \to (A \preccurlyeq D)$. Since Γ is a maximally Σ-consistent set, then $\Gamma \vdash^{\ell}_{\Sigma} (A \preccurlyeq D)$ and by theorem 3.7.4(item 1) it follows that $(A \preccurlyeq D) \in \Gamma$.

   However, this contradicts our earlier deduction that $(A \preccurlyeq D) \notin \Gamma$, and thus our assumption must be false giving us $(C \preccurlyeq D) \notin \Gamma$. Therefore, if C and D are arbitrary sentences such that $C \in \Psi_\Gamma$ and $D \notin \Psi_\Gamma$ then $(C \preccurlyeq D) \notin \Gamma$.

Thus Ψ_Γ satisfies the definition of a cut. □

Finally we prove the truth lemma, proper.

Lemma 5.3.3 (Truth Lemma)**.**
In the canonical model, our definition of v implies that we have $[\![\cdot]\!] : L_V \to \mathscr{P}(\mathscr{W}_\Sigma)$ such that for all sentences, C, the following holds

$$[\![C]\!] = \left\{ \Delta \in \mathscr{W}_\Sigma \;\middle|\; C \in \Delta \right\}.$$

In other words, we wish to show that for all maximal Σ-consistent sets, Γ

$$\Gamma \in [\![C]\!] \textit{ iff } C \in \Gamma.$$

Proof. We prove this inductively. **Base Case:**

1. If $C := \mathbb{P}_n$ for some $n \in \mathbb{N}$, then $[\![\mathbb{P}_n]\!] = v(\mathbb{P}_n) = \left\{ \Delta \in \mathscr{W}_\Sigma \;\middle|\; \mathbb{P}_n \in \Delta \right\}$.

2. If $C := \bot$, then $[\![\bot]\!] = \varnothing$ which is equal to $\left\{ \Delta \in \mathscr{W}_\Sigma \;\middle|\; \bot \in \Delta \right\}$ since $\bot$ is contained in no maximally Σ-consistent set.

Inductive Step: Next we consider the case where the sentence is more complex than a proposition or $\bot$.

Inductive Hypothesis: Suppose $[\![A]\!] = \left\{ \Delta \in \mathscr{W}_\Sigma \;\middle|\; A \in \Delta \right\}$ and $[\![B]\!] = \left\{ \Delta \in \mathscr{W}_\Sigma \;\middle|\; B \in \Delta \right\}$ as our inductive hypothesis.

We have two cases:

1. If $C := A \to B$. Let Γ be an arbitrary maximally Σ-consistent set.

$$\begin{aligned}
&\Gamma \in [\![A \to B]\!] \\
&\quad \text{iff } \Gamma \in [\![\neg A]\!] \cup [\![B]\!] \\
&\quad \text{iff } \Gamma \notin [\![A]\!] \\
&\qquad \text{or } \Gamma \in [\![B]\!] \\
&\quad \text{iff } \Gamma \notin \{\, \Delta \in \mathscr{W}_\Sigma \mid A \in \Delta \,\} \\
&\qquad \text{or } \Gamma \in \{\, \Delta \in \mathscr{W}_\Sigma \mid B \in \Delta \,\} \\
&\qquad\quad \text{(by Inductive Hypothesis)} \\
&\quad \text{iff } A \notin \Gamma \\
&\qquad \text{or } B \in \Gamma \\
&\quad \text{iff } (A \to B) \in \Gamma \\
&\qquad \text{(by theorem 3.7.4(item 4))} \\
&\quad \text{iff } \Gamma \in \{\, \Delta \in \mathscr{W}_\Sigma \mid (A \to B) \in \Delta \,\}.
\end{aligned}$$

2. If $C := A \preccurlyeq B$. Let Γ be an arbitrary maximally Σ-consistent set.

   Recall definition 2.3.3(item 4),

$$[\![A \preccurlyeq B]\!] = \{\, \Delta \in \mathscr{W}_\Sigma \mid \forall S \in \mathscr{O}(\Delta), ([\![B]\!] \bullet S \to [\![A]\!] \bullet S) \,\}.$$

   We want to prove that $\Gamma \in [\![A \preccurlyeq B]\!]$ iff $(A \preccurlyeq B) \in \Gamma$ and by the previous definition this is equivalent to proving

$$\forall S \in \mathscr{O}(\Gamma), (S \bullet [\![B]\!] \to S \bullet [\![A]\!]) \quad \text{iff} \quad (A \preccurlyeq B) \in \Gamma.$$

   We will do the proof in two directions.

   For the first direction we will prove that

$$\text{if } \exists S \in \mathscr{O}(\Gamma), (S \bullet [\![B]\!] \wedge S \not\bullet [\![A]\!]) \quad \text{then} \quad (A \preccurlyeq B) \notin \Gamma.$$

   Suppose

$$\exists S \in \mathscr{O}(\Gamma), (S \bullet [\![B]\!] \wedge S \not\bullet [\![A]\!]).$$

Let $S \in \mathscr{O}(\Gamma)$ be a sphere satisfying the sentence. Recall that by the definition of a sphere, S, is a union of a set of co-spheres (of cuts around Γ), meaning that there exists a set of cuts, $X_S \subseteq \textsc{Cuts}(\Gamma)$, such that,

$$S = \bigcup_{\Psi_\Gamma \in X_S} \mathcal{S}(\Psi_\Gamma).$$

It then follows that

$$S \bullet [\![B]\!] \quad \text{iff} \quad \exists \Psi_\Gamma \in X_S, \mathcal{S}(\Psi_\Gamma) \bullet [\![B]\!],$$

and

$$S \not\bullet [\![A]\!] \quad \text{iff} \quad \forall \Psi_\Gamma \in X_S, \mathcal{S}(\Psi_\Gamma) \not\bullet [\![A]\!].$$

Let $\Psi_\Gamma \in X_S$ be a cut whose existence is granted by the first statement. Then, for Ψ_Γ, we have

$$\mathcal{S}(\Psi_\Gamma) \bullet [\![B]\!] \quad \text{and} \quad \mathcal{S}(\Psi_\Gamma) \not\bullet [\![A]\!].$$

Applying the inductive hypothesis to each statement gives us

$$\mathcal{S}(\Psi_\Gamma) \bullet \left\{ \Delta \in \mathscr{W}_\Sigma \;\middle|\; B \in \Delta \right\} \quad \text{and} \quad \mathcal{S}(\Psi_\Gamma) \not\bullet \left\{ \Delta \in \mathscr{W}_\Sigma \;\middle|\; A \in \Delta \right\}.$$

Moreover, since $\mathcal{S}(\Psi_\Gamma)$ is a co-sphere of a cut around $\Gamma \in \mathscr{W}_\Sigma$, then by lemma 5.3.1, we obtain

$$B \notin \Psi_\Gamma \quad \text{and} \quad A \in \Psi_\Gamma.$$

Finally, by the definition of a cut we get that $(A \preccurlyeq B) \notin \Gamma$. This completes the first direction.

For the second direction we will prove that

$$\text{if } \forall S \in \mathscr{O}(\Gamma), (S \bullet [\![B]\!] \to S \bullet [\![A]\!]) \quad \text{then} \quad (A \preccurlyeq B) \in \Gamma.$$

Suppose

$$\forall S \in \mathscr{O}(\Gamma), (S \bullet [\![B]\!] \to S \bullet [\![A]\!]).$$

Let Ψ_Γ be the set of sentences:

$$\Psi_\Gamma = \{ X \mid (A \preccurlyeq X) \in \Gamma \}.$$

By lemma 5.3.2, Ψ_Γ is a cut. We have the derivation:

1. $\vdash_\Sigma A \to A$ Taut
2. $\vdash_\Sigma (A \preccurlyeq A)$ Cp

Therefore, by theorem 3.7.4(item 1), it follows that $(A \preccurlyeq A) \in \Gamma$. Furthermore, by the definition of Ψ_Γ we have that $A \in \Psi_\Gamma$.

Next, consider the co-sphere, $\mathcal{S}(\Psi_\Gamma)$, of the cut, Ψ_Γ, around Γ. Since $A \in \Psi_\Gamma$, and recalling the inductive hypothesis for A, we may apply the weak co-sphere lemma 5.3.1 to get that $\mathcal{S}(\Psi_\Gamma) \not\between [\![A]\!]$. Moreover, since $\mathcal{S}(\Psi_\Gamma) \in \mathcal{O}(\Gamma)$ then applying our assumption tells us that $\mathcal{S}(\Psi_\Gamma) \not\between [\![B]\!]$. The inductive hypothesis for B allows us to apply the weak co-sphere lemma 5.3.1 once again, which gives us that $B \in \Psi_\Gamma$. Finally, by the definition of Ψ_Γ, we get that $B \in \Psi_\Gamma$ iff $(A \preccurlyeq B) \in \Gamma$.

This concludes the second direction.

□

Now that the truth lemma has been proven we are able to obtain the general version of the co-sphere lemma.

Lemma 5.3.4 (Co-sphere Lemma (Lewis, 2001, Section 6.1))**.**
Given a canonical model $M_\Sigma = (\mathcal{W}_\Sigma, \mathcal{O}, v)$ for a V-logic, Σ, if $\mathcal{S}(\Psi_\Delta)$ is the co-sphere of a cut around a $\Delta \in \mathcal{W}_\Sigma$, then a sentence $A \in \Psi_\Delta$ iff $[\![A]\!]$ does not overlap $\mathcal{S}(\Psi_\Delta)$.

Proof. The proof follows trivially from the truth lemma 5.3.3 combined with the weak co-sphere lemma 5.3.1. □

Chapter 6

Completeness

Now that we have canonical models defined we can proceed to prove that our semantics are complete with respect to our model. That is to say that "Every valid sentence in our semantics, can be derived in our syntax." Completeness is often harder than soundness. We don't have other versions of completeness and we have to be careful in determining the criteria required for completeness to hold for a logic.

Our strategy here will be to first provide some definitions to establish the conditions required for a logic to be complete. Then we will prove the completeness theorem. Finally, we will use our definitions to obtain a slightly weaker but easier to use version of the completeness theorem.

We'll first give a few definitions

6.1 Preliminary definitions

We first give a couple useful definitions.

Definition 6.1.1 (Canonical Frame)**.**
Given the canonical model, $M_\Sigma = (\mathscr{W}_\Sigma, \mathscr{O}, \nu)$, of an axiomatic system Σ, we call the frame, $F_\Sigma = (\mathscr{W}_\Sigma, \mathscr{O})$, the **canonical frame**.

Definition 6.1.2 (Canonical Axiomatic System)**.**
An axiomatic system Σ is called a **canonical axiomatic system** if and only if

every sentence in Σ is valid over the canonical frame. In other words, Σ is a canonical axiomatic system if and only if

$$\forall A \in \Sigma, \quad F_\Sigma \vDash A.$$

6.2 Frame Definability

Definition 6.2.1 (Frame Definability)**.**
We say that a sentence A **defines** a class of frames, $\mathscr{F}$, if and only if the following condition holds

$$\forall F, \quad F \in \mathscr{F} \text{ iff } F \vDash A.$$

Example 6.2.2.
The sentence **4**: $\Box A \to \Box\Box A$ defines the class, $\mathscr{F}$, consisting of all frames that satisfy the property 4: $\forall w \in W, \forall v \in \left(\bigcup \mathscr{O}(w)\right)$,

$$\left(\bigcup \mathscr{O}(v)\right) \subseteq \left(\bigcup \mathscr{O}(w)\right).$$

Proof. We prove each direction separately. For the first direction we'll show that $\forall F$, if $F \in \mathscr{F}$ then $F \vDash$ **4**.

Let F be arbitrary and suppose that $F \in \mathscr{F}$ (and therefore satisfies property 4). Let $M = (W, \mathscr{O}, v)$ be an arbitrary model over F. Let $w \in W$ be arbitrary.

Suppose $w \in [\![\Box A]\!]$, we will show that $w \in [\![\Box\Box A]\!]$. Note that due to theorem 2.3.5(item 11), $\bigcup \mathscr{O}(w) \subseteq [\![A]\!]$. Let $v \in \bigcup \mathscr{O}(w)$ be arbitrary. Since our model satisfies property 4, then $\bigcup \mathscr{O}(v) \subseteq \bigcup \mathscr{O}(w)$. Therefore, $\bigcup \mathscr{O}(v) \subseteq [\![A]\!]$ and, by theorem 2.3.5(item 11), $v \in [\![\Box A]\!]$. Since $v \in \bigcup \mathscr{O}(w)$ was arbitrary, then it follows that $\bigcup \mathscr{O}(w) \subseteq [\![\Box A]\!]$.

Finally, by theorem 2.3.5(item 11), we have that $w \in [\![\Box\Box A]\!]$. Since $w \in W$ was arbitrary, then $[\![\Box A]\!] \subseteq [\![\Box\Box A]\!]$. Then, by lemma 2.3.6 it follows that $[\![\Box A \to \Box\Box A]\!] = W$. Since M and F were arbitrary, it follows that,

$$\forall F, \text{ if } F \in \mathscr{F} \text{ then } F \vDash \mathbf{4},$$

completing the first direction of our proof.

For the second direction we'll show that $\forall F$, if $F \notin \mathscr{F}$ then $F \nvDash \mathbf{4}$.

Suppose F is an arbitrary frame. Suppose that $F \notin \mathscr{F}$. Then F does not satisfy our property 4, meaning that

$$\exists w \in W, \exists v \in \bigcup \mathscr{O}(w), \quad \bigcup \mathscr{O}(v) \nsubseteq \bigcup \mathscr{O}(w).$$

In other words

$$\exists w \in W, \exists v \in \bigcup \mathscr{O}(w), \exists u \in W, \quad u \in \bigcup \mathscr{O}(v) \text{ and } u \notin \bigcup \mathscr{O}(w).$$

Choose, w, v, u, satisfying this sentence.

In order to show that $F \nvDash \mathbf{4}$ we need to show that there exists a world in a model over F where an instance of sentence $\mathbf{4}$ fails. Choose $M = (W, \mathscr{O}, \nu)$ to be the model over $F = (W, \mathscr{O})$ such that $\nu(\mathbb{P}_1) = \bigcup \mathscr{O}(w)$. It follows from theorem 2.3.5(item 11) that $w \in [\![\Box \mathbb{P}_1]\!]$ and by theorem 2.3.5(item 1) that $u \in [\![\neg \mathbb{P}_1]\!]$. Next, since $u \in \bigcup \mathscr{O}(v)$ and $u \in [\![\neg \mathbb{P}_1]\!]$, then, by theorem 2.3.5(item 11),

$$v \notin [\![\Box \mathbb{P}_1]\!].$$

Similarly, since $v \in \bigcup \mathscr{O}(w)$ and $v \notin [\![\Box \mathbb{P}_1]\!]$, then, by theorem 2.3.5(item 11) we obtain,

$$w \notin [\![\Box \Box \mathbb{P}_1]\!].$$

Furthermore, since $w \in [\![\Box \mathbb{P}_1]\!]$, then $w \notin \big(W - [\![\Box \mathbb{P}_1]\!]\big)$. Together, this tells us that $w \notin \big(\big(W - [\![\Box \mathbb{P}_1]\!]\big) \cup [\![\Box \Box \mathbb{P}_1]\!]\big)$. In other words, by definition 2.3.3(item 3),

$$w \notin [\![\Box \mathbb{P}_1 \to \Box \Box \mathbb{P}_1]\!].$$

Since $\Box \mathbb{P}_1 \to \Box \Box \mathbb{P}_1$ is an instance of sentence $\mathbf{4}$, then we have proven that there exists a world, w, in a model, M, over F where an instance of $\mathbf{4}$ fails. In other words, $F \nvDash \mathbf{4}$.

Hence we have that $\forall F$ if $F \notin \mathscr{F}$ then $F \nvDash \mathbf{4}$, completing our proof of the second direction. Thus our proof is complete and we have that the sentence $\mathbf{4}$ indeed defines the class consisting of all frames that satisfy the property 4. $\Box$

6.3 Frame Completeness

Definition 6.3.1 (Frame Completeness)**.**
A sentence A is said to be **frame complete with respect to a property** P if and only if for every axiomatic system Σ the following holds:

if $A \in \Sigma$ then the canonical frame, F_Σ, satisfies the property P.

Example 6.3.2 ((Lewis, 2001, Section 6.1))**.**
The sentence, $\mathbf{T} := \Box A \to A$, is frame complete with respect to the property, T:

$$\forall \Delta \in \mathscr{W}_\Sigma, \quad \Delta \in \left(\bigcup \mathscr{O}(\Delta)\right).$$

Proof. Let Σ be an arbitrary axiomatic system such that $\mathbf{T} \in \Sigma$, and let $M_\Sigma = (\mathscr{W}_\Sigma, \mathscr{O}, v)$ be the canonical model of Σ.

Let $\Delta \in \mathscr{W}_\Sigma$ be an arbitrary world. Consider the set $\Psi_\Delta = \{ B \mid \Box\neg B \in \Delta \}$. First we prove that this set is a cut around Δ.

1. Since $\Box\neg\bot$ is equivalent to $\Box\top$ which is a tautology, then $\Box\top \in \Delta$ and therefore $\bot \in \Psi_\Delta$.

2. Suppose $B \in \Psi_\Delta$ and $A \notin \Psi_\Delta$. Then $\Box\neg B \in \Delta$ and $\Box\neg A \notin \Delta$. By definition 2.1.6, $\Box\neg B$ is defined as $\bot \preccurlyeq B$. Since Δ is maximally Σ-consistent and $\Box\neg A \notin \Delta$, then we have $\neg\Box\neg A \in \Delta$. By applying definition 2.1.6 repeatedly, we get that $\neg\Box\neg A$ is equivalent to as $\neg(\bot \preccurlyeq A)$ which is equivalent to $A \prec \bot$.

   By theorem 3.3.2(item 5) we have,

   $$\vdash_\Sigma ((A \prec \bot) \wedge (\bot \preccurlyeq B)) \to (A \prec B),$$

   and therefore, by definition we have $\Delta \vdash^{\ell}_{\Sigma} (A \prec B)$. Next, by theorem 3.7.4(item 1), it follows that $(A \prec B) \in \Delta$.

Now that we've proven that it is a cut around Δ, we need to show that Δ is in the co-sphere of Ψ_Δ, $\mathscr{S}(\Psi_\Delta)$.

Since $\mathbf{T} \in \Sigma$, then by theorem 3.7.4(item 2), it follows that $\mathbf{T} \in \Delta$. Moreover, for each $B \in \Psi_\Delta$, we have $\Box\neg B \in \Delta$, then by theorem 3.7.4(item 1) it follows that $\neg B \in \Delta$. Therefore, $\neg\Psi_\Delta \subseteq \Delta$, and by definition 5.1.6, $\Delta \in \mathcal{S}(\Psi_\Delta)$.

Since Δ is inside at least one of its co-spheres then it is in the union of all of its co-spheres. Thus we have

$$\Delta \in \left(\bigcup \mathcal{O}(\Delta)\right)$$

and since $\Delta \in W$ was arbitrary, then this holds $\forall \Delta \in \mathscr{W}_\Sigma$. Finally, we observe that our property must hold over the canonical frame $F_\Sigma = (\mathscr{W}_\Sigma, \mathcal{O})$, completing the proof. □

6.4 Completeness Theorem

Theorem 6.4.1 (Completeness)**.**
Given a canonical axiomatic system, $\Sigma = \{ A_1, \dots, A_n \}$, and a family of classes of frames, $\{ \mathscr{F}_1, \dots, \mathscr{F}_n \}$, such that $\forall F$:

$$\textit{if } F \vDash A_i \textit{ then } F \in \mathscr{F}_i,$$

it must be the case that for every sentence, A,

$$\textit{if } \vDash_{\mathscr{F}} A \textit{ then } \vdash_\Sigma A.$$

where $\mathscr{F} = \bigcap \{ \mathscr{F}_1, \dots, \mathscr{F}_n \}$.

Proof. Let $\Sigma = \{ A_1, \dots, A_n \}$ be a canonical axiomatic system. Let $\{ \mathscr{F}_1, \dots, \mathscr{F}_n \}$ be a family of classes of frames such that $\forall F$, if $F \vDash A_i$ then $F \in \mathscr{F}_i$. Let $\mathscr{F} = \bigcap \{ \mathscr{F}_1, \dots, \mathscr{F}_n \}$ and

Let A be an arbitrary sentence. We will prove the contrapositive:

$$\text{If } \nvdash_\Sigma A \text{ then } \nvDash_{\mathscr{F}} A.$$

Suppose $\nvdash_\Sigma A$, then $\{ \neg A \}$ is Σ-consistent by theorem 3.6.5. By Lindenbaum's theorem 3.7.2 we can extend $\{ \neg A \}$ to a maximally Σ-consistent set Γ'. Then, consider the canonical model $M_\Sigma = (\mathscr{W}_\Sigma, \mathcal{O}, v)$ of the axiomatic system Σ. Since Γ' is a maximally Σ-consistent set, then $\Gamma' \in \mathscr{W}_\Sigma$ by definition. By the truth lemma 5.3.3, since $\{ \neg A \} \in \Gamma'$, then $\Gamma' \in [\![\neg A]\!]$. Therefore, $\Gamma' \notin [\![A]\!]$ and we get $M_\Sigma, \Gamma' \not\vDash A$.

Thus, we have a model such that $M_\Sigma \not\vDash A$, and therefore A is invalid over the canonical frame, $F_\Sigma \not\vDash A$.

It remains to show that $F_\Sigma \in \mathscr{F}$. By assumption, we know that Σ is a canonical axiomatic system, and therefore $\forall A \in \Sigma$, $F_\Sigma \vDash A$. Hence, since we have that for all frames, F, if $F \vDash A_i$ then $F \in \mathscr{F}_i$, then $F_\Sigma \in \mathscr{F}_i$ for all i such that $1 \leq i \leq n$. Therefore, F_Σ must be in the intersection of all $\mathscr{F}_i$, $F_\Sigma \in \mathscr{F}$.

This we can conclude that since $F_\Sigma \not\Vdash A$, then $\not\vDash_{\mathscr{F}} A$. □

Lemma 6.4.2.
Given a set of sentences, $\{ A_1, \dots, A_n \}$, *and a set of properties,* $\{ P_1, \dots, P_n \}$, *such that*

1. A_i *is frame complete with respect to* P_i *(see definition 6.3.1), and*
2. A_i *defines the class of frames* $F_i = \{ F \mid F \text{ satisfies property } P_i \}$ *(see definition 6.2.1).*

It follows that, $\Sigma = \{ A_1, \dots, A_n \}$, *is a canonical axiomatic system.*

Proof. Let $\{ A_1, \dots, A_n \}$ be a set of sentences and let $\{ P_1, \dots, P_n \}$ be a set of properties such that

1. A_i is frame complete with respect to P_i, and
2. A_i defines the class of frames $\mathscr{F}_i = \{ F \mid F \text{ satisfies property } P_i \}$.

Let $\Sigma = \{ A_1, \dots, A_n \}$ be an axiomatic system and let,

$$\mathscr{F} = \bigcap_{1 \leq i \leq n} \mathscr{F}_i,$$

be the intersection of all $\mathscr{F}_i$. By definition, It follows that

$$\mathscr{F} = \{ F \mid \forall i, \text{if } 1 \leq i \leq n, \text{then } F \text{ satisfies property } P_i \}.$$

Since each A_i is frame complete with respect to P_i, then the canonical frame for F_Σ satisfies every property P_i, such that $1 \leq i \leq n$. Therefore, $F_\Sigma \in \mathscr{F}$. Then, since each A_i defines a class of frames $\mathscr{F}_i$, it follows that $\mathscr{F}_i \vDash A_i$. Moreover, since $\mathscr{F} \subseteq \mathscr{F}_i$, for each i, we have that $\mathscr{F} \vDash A_i$ for each i. Therefore, for all $A \in \Sigma$, we have that $F_\Sigma \vDash A$, making Σ satisfy the definition of a canonical axiomatic system. □

This lets us formulate the completeness theorem in a slightly weaker but more familiar way.

Theorem 6.4.3 (Completeness Theorem)**.**
Given a set of sentences, $\{ A_1, \dots, A_n \}$*, and a set of properties,* $\{ P_1, \dots, P_n \}$*, such that*

1. A_i *is frame complete with respect to* P_i*, and*
2. A_i *defines the class of frames* $\mathscr{F}_i = \{ F \mid F \text{ satisfies property } P_i \}$.

It must be the case that for every sentence, A*,*

$$\text{if } \vDash_{\mathscr{F}} A \text{ then } \vdash_\Sigma A,$$

where $\Sigma = \{ A_1, \dots, A_n \}$ *and* $\mathscr{F} = \bigcap \{ \mathscr{F}_1, \dots, \mathscr{F}_n \}$.

Proof. Note that if A_i defines the class of frames $\mathscr{F}_i$, then $\forall F$, if $F \vDash A_i$ then $F \in \mathscr{F}_i$. Moreover, by lemma 6.4.2 we have that Σ is a canonical axiomatic system. So we satisfy the assumptions of theorem 6.4.1 and the result follows. □

Chapter 7

Krabbe's Counterexample

In Lewis' first edition of *Counterfactuals* Lewis incorrectly claimed, without proof, that

> "The union of the co-spheres of any given set of cuts around *i* is the co-sphere of the intersection of the cuts." (Lewis (1973))

More precisely, he incorrectly claimed that, given an axiomatic system, Σ, a maximally Σ-consistent $\Delta \in \mathscr{W}_\Sigma$, and a non-empty $X \subseteq \textsc{Cuts}(\Delta)$,

$$\bigcup_{\Psi_\Delta \in X} \mathscr{S}\left(\Psi_\Delta\right) = \mathscr{S}\left(\bigcap_{\Psi_\Delta \in X} \Psi_\Delta\right). \qquad \textbf{(False)}$$

If you take this statement as true, then spheres (as unions of co-spheres) are just co-spheres themselves. As a result, the first edition of *Counterfactuals* did not include a definition for spheres[1] and several definitions and proofs related to canonical models were simpler. For instance, in the definition of a canonical model, $\mathscr{O}(\Delta)$ was defined as the set of co-spheres of all cuts around Δ. Unfortunately, his claim not only turned out to be false but the error broke his definition of canonical models (a canonical model was itself not necessarily a sphere model under his definition).

[1]Technically, his second edition (Lewis, 2001) didn't explicitly define "spheres" either and instead just refers to them as the elements in the system of spheres.

After the first edition (Lewis, 1973) was published, Krabbe published a paper (Krabbe, 1978) demonstrating the error and recommending a remedy by altering the definition of canonical models. At the core, Krabbe's paper uses a counterexample in order to show that the union of a set of co-spheres (of a non-empty set of cuts) may not necessarily be a co-sphere itself. Krabbe's paper is extremely terse and the construction of his counterexample is non-trivial — including the bibliography, the entire paper is under three pages in length. In response to Krabbe's paper, Lewis published a second edition of *Counterfactuals* where he implemented Krabbe's recommended definition for canonical models and rewrote other definitions and proofs that were impacted.

Krabbe's attention to detail found a nuanced and counterintuitive error that directly influenced our definitions and proofs related to canonical models. Since Krabbe's paper omits a lot of technical details then we will take inspiration from him and give an explicit presentation of all of the details in his paper using the machinery we have developed.

7.1 Overview

The primary goal of Krabbe's paper is to prove Lewis' claim false. Compare Lewis' false claim with our previous result, theorem 5.2.4(item 3), and note that in order to prove that Lewis' claim is false we need to prove that our theorem does not hold when the subset relation is reversed. In other words, we need to prove that there exists an axiomatic system, Σ, a maximally Σ-consistent $\Delta \in \mathscr{W}_\Sigma$, a non-empty $X \subseteq \textsc{Cuts}(\Delta)$, and a Γ, such that

$$\Gamma \in \mathscr{S}\left(\bigcap_{\Psi_\Delta \in X} \Psi_\Delta\right) \quad \text{but} \quad \Gamma \notin \bigcup_{\Psi_\Delta \in X} \mathscr{S}\left(\Psi_\Delta\right).$$

Though this is the primary goal of Krabbe's paper, he takes his example a step further and shows that the union of said co-spheres is not the co-sphere of any cut. More precisely, he shows that there exists an axiomatic system, Σ, a maximally Σ-consistent $\Delta \in \mathscr{W}_\Sigma$, and a non-empty $X \subseteq \textsc{Cuts}(\Delta)$, such that there does not exist a cut Φ_Δ around Δ making the following true

$$\bigcup_{\Psi_\Delta \in X} \mathscr{S}\left(\Psi_\Delta\right) = \mathscr{S}(\Phi_\Delta).$$

We will split the material up into three sections:

1. The first section will contain a construction of a sphere model, K, which we'll call the Krabbe model. The goal here is to show that a set of sentences Π is satisfied by K, which we'll use to construct Δ in the next section.

2. The second section will contain definitions and constructions of Σ, $\Delta \in \mathscr{W}_\Sigma$, $X \subseteq \textsc{Cuts}(\Delta)$, and Γ.

3. The third section will contain our desired proofs.

7.2 Krabbe Model

We begin by constructing the Krabbe model.

Since Krabbe's construction is relatively involved, I have provided examples for each definition in order to help illustrate the underlying structure.

Definition 7.2.1 (Krabbe Model)**.**
First, define the following sets of natural numbers,

$$S_m^n = \{\ n + k \mid 0 \leq k < m, k \in \mathbb{N}\ \}$$
$$T_m^n = \{\ k \mid n - m \leq k, k \in \mathbb{N}\ \}.$$

Note that we can write these in interval notation as follows

$$S_m^n = [n, n + m)|_{\mathbb{N}}$$
$$T_m^n = [n - m, \infty)|_{\mathbb{N}}.$$

So, for instance, we have:

$$\begin{aligned}
S^0_m &= \{ 0, 1, \dots, m-1 \} \\
S^n_0 &= \varnothing \\
S^n_m &= \{ n, n+1, \dots, n+m-1 \} \\
T^0_m &= \{ -m, -m+1, \dots \} \\
T^n_0 &= \{ n, n+1, \dots \} \\
T^n_n &= \{ 0, 1, \dots \} = \mathbb{N} \\
T^n_m &= \{ n-m, n-m+1, \dots \} .
\end{aligned}$$

Now, we can construct the Krabbe model. Let $K = (\mathbb{N}, \mathcal{O}, v)$ where, for each $n \in \mathbb{N}$ we have,

$$\mathcal{O}(n) = \left\{ S^n_m \;\middle|\; m \in \mathbb{N} \right\} \cup \left\{ T^n_m \;\middle|\; 0 \leq m \leq n, m \in \mathbb{N} \right\},$$

and

$$v(\mathbb{P}_n) = \{ m \mid n \leq m, m \in \mathbb{N} \} .$$

For instance, the spheres around the world 2 are as follows:

$$\begin{aligned}
\mathcal{O}(2) &= \left\{ S^2_0, S^2_1, S^2_2, \dots \right\} \cup \left\{ T^2_0, T^2_1, T^2_2 \right\} \\
&= \left\{ [2,2) \;\middle|\;_{\mathbb{N}}, [2,3)|_{\mathbb{N}}, [2,4)|_{\mathbb{N}}, \dots \right\} \cup \left\{ [2,\infty) \;\middle|\;_{\mathbb{N}}, [1,\infty)|_{\mathbb{N}}, [0,\infty)|_{\mathbb{N}} \right\} \\
&= \{ \varnothing, \{ 2 \}, \{ 2,3 \}, \dots \} \cup \{ \{ 2,3,\dots \}, \{ 1,2,\dots \}, \{ 0,1 \dots \} \} .
\end{aligned}$$

Moreover, our definition of valuation gives us

$$v(\mathbb{P}_n) = [n, \infty)|_{\mathbb{N}}.$$

so, for instance, the worlds where the atomic sentence $\mathbb{P}_2$ is true are as follows:

$$v(\mathbb{P}_2) = \{ 2, 3, \dots \} .$$

Refer to the diagram in fig. 7.1. The orange intervals correspond to the $T^2_\bullet$ spheres while the red intervals correspond to the $S^2_\bullet$ spheres. The shaded area represents $v(\mathbb{P}_2)$.

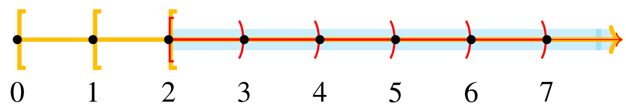


Figure 7.1: The spheres in $\mathcal{O}(2)$.

Theorem 7.2.2.
The Krabbe model is a sphere model.

Proof. We need to show that the Krabbe model is indeed a sphere model. Which means we need to show that for every $n \in \mathbb{N}$, $\mathcal{O}(n)$ is totally ordered by $\subseteq$, closed under arbitrary union, and closed under non-empty intersection.

Let $n \in \mathbb{N}$ be arbitrary. First, we'll prove closure under union and non-empty intersection. Suppose $O \subseteq \mathcal{O}(n)$ is a set of spheres around n. If O is empty, then the union is the empty set. Suppose O is non-empty and write

$$O = \{ S^n_m \mid m \in I \} \cup \{ T^n_m \mid m \in J \},$$

where I and J are index sets. Note that J must be finite due to the definition of $\mathcal{O}(n)$.

$$\bigcup O = \bigcup_{m \in I} S^n_m \cup \bigcup_{m \in J} T^n_m$$
$$\bigcap O = \bigcap_{m \in I} S^n_m \cap \bigcap_{m \in J} T^n_m.$$

If I has a smallest element, call it a_I, and if it has a largest element, call it b_I. Then we get the following:

$$I \text{ is infinite then } \bigcup_{m \in I} S^n_m = T^n_0 \quad \text{and} \quad \bigcap_{m \in I} S^n_m = S^n_{a_I}$$
$$I \text{ is finite, non-empty then } \bigcup_{m \in I} S^n_m = S^n_{b_I} \quad \text{and} \quad \bigcap_{m \in I} S^n_m = S^n_{a_I}$$
$$I \text{ is empty then } \bigcup_{m \in I} S^n_m = \varnothing \quad \text{and} \quad \bigcap_{m \in I} S^n_m = \mathbb{N}.$$

Similarly, if J has a smallest or largest element, call them a_J or b_J, respectively. Then we have the following:

$$J \text{ is finite, non-empty then } \bigcup_{m \in J} T^n_m = T^n_{b_I} \quad \text{and} \quad \bigcap_{m \in J} T^n_m = T^n_{a_I}$$
$$J \text{ is empty then } \bigcup_{m \in J} T^n_m = \varnothing \quad \text{and} \quad \bigcap_{m \in J} T^n_m = \mathbb{N}.$$

Note, that since O is non-empty, then if I is empty then J is non-empty and vice versa, so we never have $\bigcap O = \mathbb{N}$. Then, $\bigcup O$ and $\bigcap O$ simplifies to either, an $S^n_\bullet$ set, a $T^n_\bullet$ set, or a union or intersection of two such sets which will always trivially result in either an $S^n_\bullet$ set or a $T^n_\bullet$ set. Therefore, in every case we have that $\bigcap O \in \mathcal{O}(n)$ and $\bigcup O \in \mathcal{O}(n)$.

Next, we show that $\mathcal{O}(n)$ is totally ordered by $\subseteq$. Since the previous result showed us that every set in $\mathcal{O}(n)$ is either an $S^n_\bullet$ set or a $T^n_\bullet$ set then the result follows trivially, as both $S^n_\bullet$ sets and $T^n_\bullet$ are totally ordered and $S^n_\bullet$ sets are subsets of $T^n_\bullet$ sets.

Therefore, the Krabbe model, $K = (\mathbb{N}, \mathcal{O}, v)$, is a sphere model. □

Next we can introduce our special sentence Π.

Definition 7.2.3.
For each natural, $n \in \mathbb{N}$, define the sentence

$$\overline{\mathbb{P}_n} = (\mathbb{P}_0 \wedge \mathbb{P}_1 \wedge \cdots \wedge \mathbb{P}_n).$$

So, for instance, $\overline{\mathbb{P}_0} = \mathbb{P}_0$ and $\overline{\mathbb{P}_2} = (\mathbb{P}_0 \wedge \mathbb{P}_1 \wedge \mathbb{P}_2)$.

Next, define the set of sentences

$$\Pi = \left\{ \overline{\mathbb{P}_n} \prec \overline{\mathbb{P}_m} \;\middle|\; n < m \right\}.$$

So, for instance $\overline{\mathbb{P}_0} \prec \overline{\mathbb{P}_2} \in \Pi$, which is to say $(\mathbb{P}_0 \prec (\mathbb{P}_0 \wedge \mathbb{P}_1 \wedge \mathbb{P}_2)) \in \Pi$.

Theorem 7.2.4.
Π *is true at a world in a the Krabbe model.*

Proof. We'll show that Π is true at the world, 0, in the Krabbe model, $K = (\mathbb{N}, \mathcal{O}, v)$.

Recall,

$$\begin{aligned}\mathcal{O}(0) &= \{\, S^0_m \mid m \in \mathbb{N} \,\} \cup \{\, T^0_m \mid 0 \le m \le 0, m \in \mathbb{N} \,\} \\ &= \{\, S^0_m \mid m \in \mathbb{N} \,\} \cup \{\, T^0_0 \,\} \\ &= \{\, [0, m) \mid_{\mathbb{N}} | m \in \mathbb{N} \,\} \cup \{\, [0, \infty) \mid_{\mathbb{N}} \,\} \\ &= \{\, \varnothing, \{\, 0 \,\}, \{\, 0, 1 \,\}, \{\, 0, 1, 2 \,\}, \dots \,\} \cup \{\, \mathbb{N} \,\}.\end{aligned}$$

Now, recall that $v(\mathbb{P}_n) = [n, \infty)|_{\mathbb{N}}$. So,

$$\begin{aligned}v(\mathbb{P}_0) &= \mathbb{N} \\ v(\mathbb{P}_1) &= [1, \infty)|_{\mathbb{N}} \\ v(\mathbb{P}_2) &= [2, \infty)|_{\mathbb{N}} \\ &\vdots\end{aligned}$$

Visualizing this, we have fig. 7.2. The spheres are represented by the red intervals and the shaded areas represent each $v(\mathbb{P}_n)$.

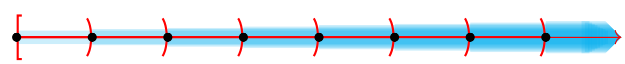


Figure 7.2: The spheres in $\mathcal{O}(0)$ with the sets given by $v(\mathbb{P}_n)$ highlighted.

Since $\overline{\mathbb{P}_n} = (\mathbb{P}_0 \wedge \mathbb{P}_1 \wedge \cdots \wedge \mathbb{P}_n)$, then using theorem 2.3.5(item 5) we can compute:

$$[\![\overline{\mathbb{P}_n}]\!] = [\![\mathbb{P}_0]\!] \cap [\![\mathbb{P}_1]\!] \cap \cdots \cap [\![\mathbb{P}_n]\!] = [\![\mathbb{P}_n]\!] = [n, \infty)|_{\mathbb{N}}.$$

In order to prove that $K, 0 \vDash \Pi$ we need to show that $0 \in [\![\Pi]\!]$ where $[\![\Pi]\!] = \bigcap_{A \in \Pi} [\![A]\!]$. In other words, we need to show that $\forall n, m \in \mathbb{N}$, such that $n < m$,

$$0 \in [\![\overline{\mathbb{P}_n} \prec \overline{\mathbb{P}_m}]\!].$$

So, let $n, m \in \mathbb{N}$, such that $n < m$ be arbitrary. Choose $S = [0, n+1)|_{\mathbb{N}}$. Note that $S \in \mathcal{O}(0)$, $[\![\mathbb{P}_n]\!] \bullet S$, and $[\![\mathbb{P}_m]\!] \not\bullet S$. Therefore,

$$\begin{aligned}
&\exists S \in \mathcal{O}(0), (([\![\mathbb{P}_n]\!] \bullet S) \wedge ([\![\mathbb{P}_m]\!] \not\bullet S)) \\
&\text{iff } \exists S \in \mathcal{O}(0), (([\![\overline{\mathbb{P}_n}]\!] \bullet S) \wedge ([\![\overline{\mathbb{P}_m}]\!] \not\bullet S)) \\
&\text{iff } 0 \in [\![\overline{\mathbb{P}_n} \prec \overline{\mathbb{P}_m}]\!] \\
&\qquad \text{(by theorem of } [\![A \prec B]\!]\text{, theorem 2.3.5(item 8)).}
\end{aligned}$$

Since, n, m such that $n < m$ were arbitrary, then this holds for all sentences in Π. Therefore, we've shown that $K, 0 \vDash \Pi$. □

Theorem 7.2.5.
Given the axiomatic system, $\mathbf{V}$*,* Π *is* $\mathbf{V}$*-consistent.*

Proof. Let $\mathcal{F}$ be the class of all frames. Then, note that since the Krabbe model, K, is a sphere model over a frame in $\mathcal{F}$ and since $K, 0 \vDash \Pi$, by theorem 7.2.4, it follows that Π is $\mathcal{F}$-satisfiable. Then note that since the set axioms of $\mathbf{V}$ are equal to the $\varnothing$ and $\mathcal{F} = \bigcap \varnothing$ we may invoke corollary 4.3.6. Therefore, Π must be $\mathbf{V}$-consistent. □

Now that we've obtained a $\mathbf{V}$-consistent set of sentences, Π, we no longer require the use of the Krabbe model.

7.3 Various Constructions

Now we will start constructing and defining the terms necessary for our counterexample, specifically Σ, $\Delta \in \mathcal{W}_\Sigma$, $X \subseteq \textsc{Cuts}(\Delta)$, and Γ. For our axiomatic system, Σ, it will be sufficient to use $\mathbf{V}$ (i.e., $\Sigma = \varnothing$).[2]

We will now begin to define $\Delta \in \mathcal{W}_{\mathbf{V}}$.

[2]Krabbe actually specifies, without proof, that his model is a **VCSU** model and provides a simplification for converting it into a **VTSA** model (see definition 2.5.2 and definition 2.5.1). He does this to show that the vast majority of the logics Lewis is interested in are affected by the counterexample, but for our purposes it's sufficient to use **V**.

Definition 7.3.1.
Since Π is **V**-consistent (see theorem 7.2.5), then we can obtain maximally **V**-consistent supersets of Π by theorem 3.7.2. Choose Δ to be an arbitrary maximally **V**-consistent superset of Π. Note that since Δ is maximally **V**-consistent, then $\Delta \in \mathscr{W}_{\mathbf{V}}$.

Next, we will begin to define $X \subseteq \textsc{Cuts}(\Delta)$. First, we'll define a family of cuts, then we'll prove that they are indeed cuts.

Definition 7.3.2.
Given the axiomatic system, **V**. By definition 7.3.1, we have the maximally **V**-consistent set, Δ. For each $n \in \mathbb{N}$ we define a set of sentences Ψ_n as follows:

$$\Psi_n = \left\{ A \,\middle|\, \left(\overline{\mathbb{P}_n} \preccurlyeq A\right) \in \Delta \right\}.$$

Moreover, define $X = \left\{ \Psi_n \,\middle|\, n \in \mathbb{N} \right\}$.

We need to prove that each Ψ_n is indeed a cut around Δ, so we do that now.

Theorem 7.3.3.
For each $n \in \mathbb{N}$, Ψ_n, as defined in definition 7.3.2, is a cut around Δ.

Proof. Let $n \in \mathbb{N}$ be arbitrary and let Ψ_n be defined as in definition 7.3.2.

We will prove Ψ_n satisfies the definition of a cut definition 5.1.1.

First, note that we have the derivation:

1. $\vdash_{\mathbf{V}} \bot \to \overline{\mathbb{P}_n}$ Taut
2. $\vdash_{\mathbf{V}} \overline{\mathbb{P}_n} \preccurlyeq \bot$ Cp,1

Since this is a derivation then, by theorem 3.7.4(item 1), it must be in Δ. Therefore, by definition 7.3.2, $\bot \in \Psi_n$.

Next, suppose $B \in \Psi_n$ but $A \notin \Psi_n$. By, definition 7.3.2, that means $\left(\overline{\mathbb{P}_n} \preccurlyeq B\right) \in \Delta$ and $\left(\overline{\mathbb{P}_n} \preccurlyeq A\right) \notin \Delta$. Since Δ is maximally **V**-consistent, we obtain $\left(A \prec \overline{\mathbb{P}_n}\right) \in \Delta$. Moreover, by theorem 3.3.2(item 5) we have the theorem

$$\vdash_{\mathbf{V}} \left(\left(A \prec \overline{\mathbb{P}_n}\right) \wedge \left(\overline{\mathbb{P}_n} \preccurlyeq B\right)\right) \to (A \prec B).$$

By, definition 3.5.1, we get $\Delta \vdash^{\ell}_{\mathbf{V}} (A \prec B)$ and by theorem 3.7.4(item 1) we have $(A \prec B) \in \Delta$.

Therefore, by definition 5.1.1, it follows that Ψ_n is a cut around Δ and since $n \in \mathbb{N}$ was arbitrary we have that for all $n \in \mathbb{N}$, Ψ_n is a cut around Δ. □

With this we have our family of cuts, $X \subseteq \textsc{Cuts}(\Delta)$ properly defined.

Our next goal is to define Γ. However, first it will be helpful to describe the intersection of X and to do that we will need to prove this lemma.

Lemma 7.3.4.
Given Δ as defined in definition 7.3.1, for any sentence, A,

$$\forall n \in \mathbb{N}, \left(\overline{\mathbb{P}_n} \preccurlyeq A\right) \in \Delta \qquad \textit{iff} \qquad \forall n \in \mathbb{N}, \left(\overline{\mathbb{P}_n} \prec A\right) \in \Delta.$$

Proof. Let Δ be defined in definition 7.3.1. Let A be a sentence.

We prove each direction separately.

For the first direction, suppose $\forall n \in \mathbb{N}, \left(\overline{\mathbb{P}_n} \preccurlyeq A\right) \in \Delta$. Let $m \in \mathbb{N}$ be arbitrary. We want to show that $\forall n \in \mathbb{N}, \left(\overline{\mathbb{P}_n} \prec A\right) \in \Delta$. So, let $n \in \mathbb{N}$ be arbitrary, such that $m < n$. Due to our assumption, we obtain that $\left(\overline{\mathbb{P}_n} \preccurlyeq A\right) \in \Delta$. Recall that by definition 7.2.3, $\left(\overline{\mathbb{P}_m} \prec \overline{\mathbb{P}_n}\right) \in \Pi$, and since $\Pi \subseteq \Delta$, by definition 7.3.1, then $\left(\overline{\mathbb{P}_m} \prec \overline{\mathbb{P}_n}\right) \in \Delta$.

Furthermore, by, theorem 3.3.2(item 5), it follows that,

$$\vdash_{\mathbf{V}} \left[\left(\overline{\mathbb{P}_m} \prec \overline{\mathbb{P}_n}\right) \wedge \left(\overline{\mathbb{P}_n} \preccurlyeq A\right)\right] \rightarrow \left(\overline{\mathbb{P}_m} \prec A\right),$$

is a theorem and since $\left(\overline{\mathbb{P}_m} \prec \overline{\mathbb{P}_n}\right) \in \Delta$ and $\left(\overline{\mathbb{P}_n} \preccurlyeq A\right) \in \Delta$, then by definition 3.5.1 we have

$$\Delta \vdash^{\ell}_{\mathbf{V}} \left(\overline{\mathbb{P}_m} \prec A\right).$$

Moreover, since Δ is maximally $\mathbf{V}$-consistent, then by, theorem 3.7.4(item 1), we get that $\left(\overline{\mathbb{P}_m} \prec A\right) \in \Delta$. Finally, since $m \in \mathbb{N}$ was arbitrary, then we have

$$\forall n \in \mathbb{N}, \left(\overline{\mathbb{P}_n} \prec A\right) \in \Delta.$$

This completes the first direction.

For the second direction, suppose $\forall n \in \mathbb{N}, \left(\overline{\mathbb{P}_n} \prec A\right) \in \Delta$. Let $n \in \mathbb{N}$ be arbitrary. Then $\left(\overline{\mathbb{P}_n} \prec A\right) \in \Delta$. Recall the theorem 3.3.2(item 3),

$$\vdash_{\mathbf{V}} \left(\overline{\mathbb{P}_n} \prec A\right) \rightarrow \left(\overline{\mathbb{P}_n} \preccurlyeq A\right).$$

Since $\left(\overline{\mathbb{P}_n} \prec A\right) \in \Delta$ then we have

$$\Delta \vdash^{\ell}_{\mathbf{V}} \left(\overline{\mathbb{P}_n} \preccurlyeq A\right).$$

Since Δ is maximally $\mathbf{V}$-consistent, then by, theorem 3.7.4(item 1), we get that $\left(\overline{\mathbb{P}_n} \preccurlyeq A\right) \in \Delta$.

Since $n \in \mathbb{N}$ was arbitrary, then we have

$$\forall n \in \mathbb{N}, \left(\overline{\mathbb{P}_n} \preccurlyeq A\right) \in \Delta.$$

This completes the second direction. □

Now we can describe the intersection of X.

Theorem 7.3.5.
Given Ψ_n and X as defined in definition 7.3.2,

$$\bigcap_{\Psi_n \in X} \Psi_n = \left\{ A \;\middle|\; \forall n \in \mathbb{N}, \left(\overline{\mathbb{P}_n} \prec A\right) \in \Delta \right\}.$$

Proof. Let Ψ_n and X be defined as in definition 7.3.2

$$A \in \bigcap_{\Psi_n \in X} \Psi_n$$

iff $\forall n \in \mathbb{N}, A \in \Psi_n$

iff $\forall n \in \mathbb{N}, \left(\overline{\mathbb{P}_n} \preccurlyeq A\right) \in \Delta$

(by definition of Ψ_n, definition 7.3.2)

iff $\forall n \in \mathbb{N}, \left(\overline{\mathbb{P}_n} \prec A\right) \in \Delta$

(by lemma 7.3.4)

iff $A \in \left\{ A \;\middle|\; \forall n \in \mathbb{N}, \left(\overline{\mathbb{P}_n} \prec A\right) \in \Delta \right\}$. □

Moreover, by theorem 5.2.3(item 3), $\bigcap_{\Psi_n \in X} \Psi_n$ is itself a cut.

Finally, we can begin working directly toward defining Γ. We'll do this by first defining Θ, showing that Θ is **V**-consistent, and finally defining Γ as a maximally **V**-consistent superset of Θ.

Definition 7.3.6.
Given $\overline{\mathbb{P}_n}$ defined as in definition 7.2.3 and Ψ_n and X as defined in definition 7.3.2, define

$$\Theta = \left\{ \overline{\mathbb{P}_n} \;\middle|\; n \in \mathbb{N} \right\} \cup \neg \left(\bigcap_{\Psi_n \in X} \Psi_n \right).$$

Unlike our strategy with Π, we are able to show that Θ is **V**-consistent without having to construct a model.

Theorem 7.3.7.
*Θ, as defined by definition 7.3.6, is **V**-consistent.*

Proof. Suppose for the sake of contradiction that Θ is **V**-inconsistent. Then we have the derivation

$$\Theta \vdash^{\ell}_{\mathbf{V}} \bot,$$

such that for some sentences $\{ A_1, A_2, \dots, A_m \} \subseteq \bigcap_{\Psi_n \in X} \Psi_n$ and some sentences $\{ B_1, B_2 \dots, B_j \} \subseteq \left\{ \overline{\mathbb{P}_n} \;\middle|\; n \in \mathbb{N} \right\}$:

$$\vdash_{\mathbf{V}} \left(\neg A_1 \wedge \neg A_2 \wedge \cdots \wedge \neg A_m \wedge B_1 \wedge B_2 \wedge \cdots \wedge B_j \right) \to \bot.$$

However, due to the properties of conjunction and by definition 7.2.3, there exists some $n \in \mathbb{N}$ such that we have the following derivation:

1.	$\vdash_{\mathbf{V}} \left[(\neg A_1 \wedge \neg A_2 \wedge \cdots \wedge \neg A_m) \wedge (B_1 \wedge B_2 \wedge \cdots \wedge B_j) \right] \to \bot$	Assumption
2.	$\vdash_{\mathbf{V}} (B_1 \wedge B_2 \wedge \cdots \wedge B_j) \leftrightarrow (B_1 \wedge B_2 \wedge \cdots \wedge B_j)$	Taut
3.	$\vdash_{\mathbf{V}} (B_1 \wedge B_2 \wedge \cdots \wedge B_j) \leftrightarrow \overline{\mathbb{P}_n}$	Pl,2
4.	$\vdash_{\mathbf{V}} \left(\neg A_1 \wedge \neg A_2 \wedge \cdots \wedge \neg A_m \wedge \overline{\mathbb{P}_n} \right) \to \bot$	Eq,1,3

But by definition, that would mean that the following set is **V**-inconsistent,

$$\left\{ \overline{\mathbb{P}_n} \right\} \cup \neg \left(\bigcap_{\Psi_n \in X} \Psi_n \right).$$

However, by theorem 5.1.5(item 5), we know that this set is **V**-consistent. Therefore, by contradiction, Θ must be **V**-consistent as well. □

Finally, we are able define Γ.

Definition 7.3.8.
Since Θ is **V**-consistent, then we can obtain maximally **V**-consistent supersets, by theorem 3.7.2. Choose Γ to be an arbitrary maximally **V**-consistent superset of Θ.

As a quick summary, we have given definitions and constructions of $\Sigma = \mathbf{V}$, $\Delta \in \mathscr{W}_\Sigma$ (definition 7.3.1), $X \subseteq \textsc{Cuts}(\Delta)$ (definition 7.3.2), and Γ (definition 7.3.8).

7.4 Krabbe's Main Results

We can finally prove our two main results. For the first one, we'll prove that Lewis' claim,

> "The union of the co-spheres of any given set of cuts around *i* is the co-sphere of the intersection of the cuts," (Lewis (1973))

is false.

For cleanliness we will split the argument in two parts and prove each part separately. First we show that, given the definitions and constructions we developed in the last section, Γ is contained in the co-sphere of the union of our set of cuts.

Lemma 7.4.1.
Given the axiomatic system, $\mathbf{V}$*,* Δ *as defined by definition 7.3.1,* X *as defined by definition 7.3.2, and* Γ *as defined by definition 7.3.8, it follows that*

$$\Gamma \in \mathcal{S}\left(\bigcap_{\Psi_n \in X} \Psi_n\right).$$

Proof. Let Σ to be the axiomatic system $\mathbf{V}$ (i.e., $\Sigma = \varnothing$), Δ as defined by definition 7.3.1, X as defined by definition 7.3.2, and Γ as defined by definition 7.3.8.

Since $\neg\left(\bigcap_{\Psi_n \in X} \Psi_n\right) \subseteq \Theta$ by definition 7.3.6, then by definition 7.3.8 it follows that $\neg\left(\bigcap_{\Psi_n \in X} \Psi_n\right) \subseteq \Gamma$. Therefore, by definition 5.1.6, we have $\Gamma \in \mathcal{S}\left(\bigcap_{\Psi_n \in X} \Psi_n\right)$. □

Next we show that, given the definitions and constructions we developed in the last section, Γ is not contained in the union of the co-sphere of our set of cuts.

Lemma 7.4.2.
Given the axiomatic system, $\mathbf{V}$*,* Δ *as defined by definition 7.3.1,* X *as defined by definition 7.3.2, and* Γ *as defined by definition 7.3.8, it follows that*

$$\Gamma \notin \bigcup_{\Psi_n \in X} \mathcal{S}(\Psi_n).$$

Proof. Let Σ to be the axiomatic system $\mathbf{V}$ (i.e., $\Sigma = \varnothing$), Δ as defined by definition 7.3.1, X as defined by definition 7.3.2, and Γ as defined by definition 7.3.8.

Let $n \in \mathbb{N}$ be arbitrary. Since $\vdash_{\mathbf{V}} \left(\overline{\mathbb{P}_n} \preccurlyeq \overline{\mathbb{P}_n}\right)$ is a theorem, then by theorem 3.7.4(item 1), $\overline{\mathbb{P}_n} \in \Psi_n$. Therefore, $\neg\overline{\mathbb{P}_n} \in \neg\Psi_n$. Since, $\mathcal{S}(\Psi_n)$, is the set of maximally $\mathbf{V}$-consistent supersets of $\neg\Psi_n$, then no maximally $\mathbf{V}$-consistent set in the co-sphere, $\mathcal{S}(\Psi_n)$, can contain $\overline{\mathbb{P}_n}$.

Suppose for the sake of contradiction that $\Gamma \in \bigcup_{\Psi_n \in X} \mathcal{S}(\Psi_n)$. Then there exists some $k \in \mathbb{N}$, such that, $\Gamma \in \mathcal{S}(\Psi_k)$ and therefore, since no set in $\mathcal{S}(\Psi_k)$ can contain $\overline{\mathbb{P}_k}$, it follows that $\overline{\mathbb{P}_k} \notin \Gamma$. However, since $\left\{ \overline{\mathbb{P}_n} \;\middle|\; n \in \mathbb{N} \right\} \subseteq \Gamma$, by definition, then we have a contradiction. Thus we conclude that $\Gamma \notin \bigcup_{\Psi_n \in X} \mathcal{S}(\Psi_n)$. □

By combining our two previous results we obtain our first main result: A proof showing that Lewis' claim is false.

Theorem 7.4.3.
There exists an axiomatic system, Σ, a maximally Σ-consistent $\Delta \in \mathscr{W}_\Sigma$, a non-empty $X \subseteq$ Cuts(Δ), and a Γ, such that

$$\Gamma \in \mathcal{S}\left(\bigcap_{\Psi_\Delta \in X} \Psi_\Delta\right) \quad \textit{but} \quad \Gamma \notin \bigcup_{\Psi_\Delta \in X} \mathcal{S}\left(\Psi_\Delta\right).$$

Proof. Choose Σ to be the axiomatic system $\mathbf{V}$ (i.e., $\Sigma = \varnothing$), Δ as defined by definition 7.3.1, X as defined by definition 7.3.2, and Γ as defined by definition 7.3.8.

Then, by lemma 7.4.1 and lemma 7.4.2 the result follows. □

Before moving onto our second main result we will prove a useful corollary.

Corollary 7.4.4.
Given the axiomatic system, $\mathbf{V}$, Δ as defined by definition 7.3.1, X as defined by definition 7.3.2, and Γ as defined by definition 7.3.8, it follows that for any sentence, B,

$$\textit{if } B \in \Gamma \textit{ then } B \notin \bigcap_{\Psi_n \in X} \Psi_n.$$

Proof. Let Σ to be the axiomatic system $\mathbf{V}$ (i.e., $\Sigma = \varnothing$), Δ as defined by definition 7.3.1, X as defined by definition 7.3.2, and Γ as defined by definition 7.3.8. Let B be an arbitrary sentence. Suppose $B \in \Gamma$. Suppose for the sake of contradiction that $B \in \bigcap_{\Psi_n \in X} \Psi_n$

By lemma 7.4.1, $\Gamma \in \mathcal{S}\left(\bigcap_{\Psi_n \in X} \Psi_n\right)$. Then, by the definition of co-sphere, definition 5.1.6, $\neg\left(\bigcap_{\Psi_n \in X} \Psi_n\right) \subseteq \Gamma$.

Since $B \in \left(\bigcap_{\Psi_n \in X} \Psi_n\right)$ then $\neg B \in \Gamma$ and thus since Γ is maximally $\mathbf{V}$-consistent, then $B \notin \Gamma$. This is a contradiction, and therefore we have that if $B \in \Gamma$ then $B \notin \bigcap_{\Psi_n \in X} \Psi_n$. □

We've finally reached our second and final main result. In Krabbe's counterexample, the union of co-spheres of our set of cuts does not yield the co-sphere of any cut.

Theorem 7.4.5.
There exists an axiomatic system, Σ, a maximally Σ-consistent $\Delta \in \mathscr{W}_\Sigma$, and a non-empty $X \subseteq$ Cuts(Δ), such that there does not exist a cut Φ_Δ around Δ making the following true

$$\bigcup_{\Psi_\Delta \in X} \mathcal{S}\left(\Psi_\Delta\right) = \mathcal{S}(\Phi_\Delta).$$

Proof. Choose Σ to be the axiomatic system $\mathbf{V}$ (i.e., $\Sigma = \varnothing$), Δ as defined by definition 7.3.1, and X as defined by definition 7.3.2.

Suppose, for the sake of contradiction that there exists a cut, Φ_Δ, such that $\bigcup_{\Psi_n \in X} \mathcal{S}(\Psi_n) = \mathcal{S}(\Phi_\Delta)$.

By lemma 7.4.2, we have $\Gamma \notin \bigcup_{\Psi_n \in X} \mathcal{S}(\Psi_n)$, where Γ is defined as in definition 7.3.8. Then,

$$\Gamma \notin \bigcup_{\Psi_n \in X} \mathcal{S}(\Psi_n)$$

iff $\Gamma \notin \mathcal{S}(\Phi_\Delta)$

(by definition of $\mathcal{S}(\Phi_\Delta)$, definition 5.1.6)

iff $\neg(\Phi_\Delta) \nsubseteq \Gamma$

iff $\exists B, \neg B \in \neg(\Phi_\Delta)$ and $\neg B \notin \Gamma$

iff $\exists B, B \in \Phi_\Delta$ and $B \in \Gamma$

(by definition of maximally **V**-consistent, definition 3.7.1)

only if $\exists B, B \in \Phi_\Delta$ and $B \notin \bigcap_{\Psi_n \in X} \Psi_n$

(by corollary 7.4.4)

iff $\exists B, B \in \Phi_\Delta$ and $B \notin \left\{ A \;\middle|\; \forall n \in \mathbb{N}, \left(\overline{\mathbb{P}_n} \prec A\right) \in \Delta \right\}$

(by theorem 7.3.5)

iff $\exists B, B \in \Phi_\Delta$ and $\exists k \in \mathbb{N}, \left(\overline{\mathbb{P}_k} \prec B\right) \notin \Delta$

only if $\exists k \in \mathbb{N}, \overline{\mathbb{P}_k} \in \Phi_\Delta$

(by theorem 5.1.5(item 2)).

However, we will now show that $\forall n \in \mathbb{N}, \overline{\mathbb{P}_n} \notin \Phi_\Delta$, in order to obtain a contradiction.

Let $j \in \mathbb{N}$ be arbitrary and define the set of sentences:

$$\Xi = \left\{ \overline{\mathbb{P}_j} \right\} \cup \neg\Psi_{j+1}.$$

By theorem 5.1.5(item 5), Ξ is **V**-consistent. Let Ω be a maximal **V**-consistent extension of Ξ by theorem 3.7.2. Since $\neg\Psi_{j+1} \subseteq \Omega$, then $\Omega \in \mathcal{S}(\Psi_{j+1})$, by definition 5.1.6. Then, it follows that $\Omega \in \bigcup_{\Psi_n \in X} \mathcal{S}(\Psi_n)$ and since we assumed

$$\bigcup_{\Psi_n \in X} \mathcal{S}(\Psi_n) = \mathcal{S}(\Phi_\Delta),$$

then we also have that $\Omega \in \mathcal{S}(\Phi_\Delta)$ and thus $\neg\Phi_\Delta \subseteq \Omega$.

Moreover, $\overline{\mathbb{P}_j} \in \Omega$ by construction. However, that means that $\overline{\mathbb{P}_j} \notin \Phi_\Delta$. If it were, then $\neg\overline{\mathbb{P}_j}$ would be in $\neg\Phi_\Delta$ and therefore also in $\neg\overline{\mathbb{P}_j} \in \Omega$. Then, given that Ω is maximally **V**-consistent we would obtain a contradiction.

Since $j \in \mathbb{N}$ was arbitrary, then it follows that $\forall n \in \mathbb{N}, \overline{\mathbb{P}_n} \notin \Phi_\Delta$. This gives us a contradiction with our earlier result and therefore it must be the case that there does not exist a cut, Φ_Δ, such that $\bigcup_{\Psi_n \in X} \mathcal{S}(\Psi_n) = \mathcal{S}(\Phi_\Delta)$. □

It is because of these results that one cannot define co-spheres as spheres as Lewis (1973) had done in the first edition of Counterfactuals. In his paper, Krabbe suggests defining spheres as unions of sets of co-spheres and that is the solution adopted both in the second edition of*Counterfactuals* (Lewis, 2001) and in this thesis.

Chapter 8

Additional Proofs

Some groups of theorems result in a large collection of very similar proofs. In those cases, it is preferable to move the proofs out of the way so that they don't break up the flow of the text. This chapter contains all of those additional proofs.

8.1 Generalized Valuation Proofs

We prove the remaining proofs for theorem 2.3.5 here.

We will use the following lemma to save some steps in a few proofs.

Lemma 8.1.1.
Given a model with a generalized valuation, a sentence A, and a set $S \subseteq W$. The following holds

$$[\![\neg A]\!] \not\between S \quad \textit{iff} \quad S \subseteq [\![A]\!].$$

Proof. First note that,

$$[\![\neg A]\!] \not\between S \text{ iff } [\![\neg A]\!] \cap S = \varnothing$$
$$\text{(by definition of } \not\between\text{, definition 2.3.1).}$$

Next, let $w \in S$ be arbitrary. Then $w \in W$, since $S \subseteq W$. Then, $w \notin [\![\neg A]\!]$ because otherwise $w \in [\![\neg A]\!] \cap S$ which is a contradiction. Therefore, $w \in [\![A]\!]$, and hence $S \subseteq [\![A]\!]$. ☐

Theorem 8.1.2.
In a sphere model, $M = (W, \mathcal{O}, v)$, the following also hold for any sentences A and B:

1. $[\![\neg A]\!] = W - [\![A]\!]$
2. $[\![\neg\neg A]\!] = [\![A]\!]$
3. $[\![A \to B]\!] = [\![\neg A]\!] \cup [\![B]\!]$
4. $[\![\top]\!] = W$
5. $[\![A \wedge B]\!] = [\![A]\!] \cap [\![B]\!]$
6. $[\![A \vee B]\!] = [\![A]\!] \cup [\![B]\!]$
7. $[\![A \leftrightarrow B]\!] = ([\![A]\!] \cap [\![B]\!]) \cup ([\![\neg A]\!] \cap [\![\neg B]\!])$
8. $[\![A \prec B]\!] = \{\ w \in W \mid \exists S \in \mathcal{O}(w), ([\![A]\!] \bullet S \wedge [\![B]\!] \not\bullet S)\ \}$
9. $[\![A \approx B]\!] = \{\ w \in W \mid \forall S \in \mathcal{O}(w), ([\![B]\!] \bullet S \leftrightarrow [\![A]\!] \bullet S)\ \}$
10. $[\![\Diamond A]\!] = \left\{\ w \in W \;\middle|\; [\![A]\!] \bullet \bigcup \mathcal{O}(w)\ \right\}$
11. $[\![\Box A]\!] = \left\{\ w \in W \;\middle|\; \bigcup \mathcal{O}(w) \subseteq [\![A]\!]\ \right\}$
12. $[\![\Diamond\!\!\cdot\ A]\!] = \{\ w \in W \mid \forall S \in \mathcal{O}(w), S \neq \varnothing \to [\![A]\!] \bullet S\ \}$
13. $[\![\boxdot A]\!] = \{\ w \in W \mid \exists S \in \mathcal{O}(w), S \neq \varnothing \wedge S \subseteq [\![A]\!]\ \}$
14. $[\![A \mathrel{\Box\!\!\Rightarrow} B]\!] = \left\{\ w \in W \;\middle|\; \begin{array}{l} \exists S \in \mathcal{O}(w), \\ \quad (([\![A]\!] \bullet S) \wedge (([\![A]\!] \cap S) \subseteq [\![B]\!])) \end{array} \right\}$
15. $[\![A \mathrel{\Diamond\!\!\Rightarrow} B]\!] = \left\{\ w \in W \;\middle|\; \begin{array}{l} \forall S \in \mathcal{O}(w), \\ \quad ([\![A]\!] \bullet S \to ([\![A]\!] \cap [\![B]\!]) \bullet S) \end{array} \right\}$
16. $[\![A \mathrel{\Box\!\!\rightarrow} B]\!] = \left\{\ w \in W \;\middle|\; \begin{array}{l} [\![A]\!] \bullet \bigcup \mathcal{O}(w) \to \exists S \in \mathcal{O}(w), \\ \quad (([\![A]\!] \bullet S) \wedge (([\![A]\!] \cap S) \subseteq [\![B]\!])) \end{array} \right\}$
17. $[\![A \mathrel{\Diamond\!\!\rightarrow} B]\!] = \left\{\ w \in W \;\middle|\; \begin{array}{l} [\![A]\!] \bullet \bigcup \mathcal{O}(w) \wedge \forall S \in \mathcal{O}(w), \\ \quad ([\![A]\!] \bullet S \to ([\![A]\!] \cap [\![B]\!]) \bullet S) \end{array} \right\}$

Proof. Let $M = (W, \mathcal{O}, v)$ be a sphere model and let A and B be arbitrary sentences. Let $w \in W$ be arbitrary[1]

We prove each statement.

2.

$$
\begin{aligned}
&[\![\neg\neg A]\!] \\
&\quad = W - (W - [\![A]\!]) \\
&\qquad \text{(by theorem of } [\![\neg A]\!], \text{ theorem 2.3.5(item 1))} \\
&\quad = [\![A]\!].
\end{aligned}
$$

3.

$$
\begin{aligned}
&[\![A \to B]\!] \\
&\quad = (W - [\![A]\!]) \cup [\![B]\!] \\
&\qquad \text{(by definition of } [\![A \to B]\!], \text{ definition 2.3.3(item 3))} \\
&\quad = [\![\neg A]\!] \cup [\![B]\!] \\
&\qquad \text{(by theorem of } [\![\neg A]\!], \text{ theorem 2.3.5(item 1)).}
\end{aligned}
$$

4.

$$
\begin{aligned}
&[\![\top]\!] \\
&\quad = [\![\neg\bot]\!] \\
&\qquad \text{(by definition of } \top, \text{ definition 2.1.6)} \\
&\quad = W - [\![\bot]\!] \\
&\qquad \text{(by theorem of } [\![\neg A]\!], \text{ theorem 2.3.5(item 1))} \\
&\quad = W - \varnothing \\
&\qquad \text{(by definition of } [\![\bot]\!], \text{ definition 2.3.3(item 2))} \\
&\quad = W.
\end{aligned}
$$

[1]Since the domain of generalized valuation is subsets of W, then this assumption doesn't weaken our proofs. It will, however, streamline our proofs by allowing us to write $w \notin [\![A]\!]$ instead of $w \in (W - [\![A]\!])$ and letting us use statements like $w \in \{v \in W \mid \phi(v)\}$ iff $\phi(w)$.

5.

$$
\begin{aligned}
&[\![A \wedge B]\!] \\
&\quad = [\![\neg(A \to \neg B)]\!] \\
&\qquad \text{(by definition of } A \wedge B\text{, definition 2.1.6)} \\
&\quad = W - [\![(A \to \neg B)]\!] \\
&\qquad \text{(by theorem of } [\![\neg A]\!]\text{, theorem 2.3.5(item 1))} \\
&\quad = W - ([\![\neg A]\!] \cup [\![\neg B]\!]) \\
&\qquad \text{(by theorem of } [\![A \to B]\!]\text{, theorem 2.3.5(item 3))} \\
&\quad = W - [(W - [\![A]\!]) \cup (W - [\![B]\!])] \\
&\qquad \text{(by theorem of } [\![\neg A]\!]\text{, theorem 2.3.5(item 1))} \\
&\quad = W - [W - ([\![A]\!] \cap [\![B]\!])] \\
&\quad = [\![A]\!] \cap [\![B]\!].
\end{aligned}
$$

6.

$$
\begin{aligned}
&[\![A \vee B]\!] \\
&\quad = [\![\neg A \to B]\!] \\
&\qquad \text{(by definition of } A \vee B\text{, definition 2.1.6)} \\
&\quad = [\![\neg\neg A]\!] \cup [\![B]\!] \\
&\qquad \text{(by theorem of } [\![\neg\neg A]\!]\text{, theorem 2.3.5(item 2))} \\
&\quad = [\![A]\!] \cup [\![B]\!].
\end{aligned}
$$

7.

$$
\begin{aligned}
&[\![A \leftrightarrow B]\!] \\
&\quad = [\![((A \to B) \wedge (A \leftarrow B))]\!] \\
&\qquad \text{(by definition of } A \leftrightarrow B\text{, definition 2.1.6)} \\
&\quad = [\![A \to B]\!] \cap [\![A \leftarrow B]\!] \\
&\qquad \text{(by theorem of } [\![A \wedge B]\!]\text{, theorem 2.3.5(item 5))} \\
&\quad = ([\![\neg A]\!] \cup [\![B]\!]) \cap ([\![\neg B]\!] \cup [\![A]\!]) \\
&\qquad \text{(by theorem of } [\![A \to B]\!]\text{, theorem 2.3.5(item 3))} \\
&\quad = ([\![A]\!] \cap [\![B]\!]) \cup ([\![\neg A]\!] \cap [\![\neg B]\!]).
\end{aligned}
$$

8.

$$
\begin{aligned}
&w \in [\![A \prec B]\!] \\
&\quad \text{iff } w \in [\![\neg(B \preccurlyeq A)]\!] \\
&\qquad \text{(by definition of } A \prec B\text{, definition 2.1.6)} \\
&\quad \text{iff } w \notin [\![B \preccurlyeq A]\!] \\
&\qquad \text{(by theorem of } [\![\neg A]\!]\text{, theorem 2.3.5(item 1))} \\
&\quad \text{iff } w \notin \{\, v \in W \mid \forall S \in \mathcal{O}(v), ([\![A]\!] \bullet S \to [\![B]\!] \bullet S) \,\} \\
&\qquad \text{(by definition of } [\![A \preccurlyeq B]\!]\text{, definition 2.3.3(item 4))} \\
&\quad \text{iff not } \forall S \in \mathcal{O}(w), ([\![A]\!] \bullet S \to [\![B]\!] \bullet S) \\
&\quad \text{iff } \exists S \in \mathcal{O}(w), ([\![A]\!] \bullet S \wedge [\![B]\!] \not\bullet S) \\
&\quad \text{iff } w \in \{\, v \in W \mid \exists S \in \mathcal{O}(v), ([\![A]\!] \bullet S \wedge [\![B]\!] \not\bullet S) \,\}.
\end{aligned}
$$

9.

$w \in [\![A \approx B]\!]$

iff $w \in [\![((A \preccurlyeq B) \wedge (B \preccurlyeq A))]\!]$

(by definition of $A \approx B$, definition 2.1.6)

iff $w \in [\![(A \preccurlyeq B)]\!]$ and $w \in [\![(B \preccurlyeq A)]\!]$

iff $w \in \{ v \in W \mid \forall S \in \mathcal{O}(v), ([\![B]\!] \bullet S \rightarrow [\![A]\!] \bullet S) \}$

and $w \in \{ v \in W \mid \forall S \in \mathcal{O}(v), ([\![A]\!] \bullet S \rightarrow [\![B]\!] \bullet S) \}$

(by definition of $[\![A \preccurlyeq B]\!]$, definition 2.3.3(item 4))

iff $\forall S \in \mathcal{O}(w), ([\![B]\!] \bullet S \rightarrow [\![A]\!] \bullet S)$

and $\forall S \in \mathcal{O}(w), ([\![A]\!] \bullet S \rightarrow [\![B]\!] \bullet S)$

iff $\forall S \in \mathcal{O}(w), ([\![B]\!] \bullet S \leftrightarrow [\![A]\!] \bullet S)$

iff $w \in \{ v \in W \mid \forall S \in \mathcal{O}(v), ([\![B]\!] \bullet S \leftrightarrow [\![A]\!] \bullet S) \}$

(by definition of $[\![A \preccurlyeq B]\!]$, definition 2.3.3(item 4)).

10.

$w \in [\![\Diamond A]\!]$

iff $w \in [\![A \prec \bot]\!]$

(by definition of $\Diamond A$, definition 2.1.6)

iff $w \in \{ v \in W \mid \exists S \in \mathcal{O}(v), ([\![A]\!] \bullet S \wedge [\![\bot]\!] \not\bullet S) \}$

(by theorem of $[\![A \prec B]\!]$, theorem 2.3.5(item 8))

iff $w \in \{ v \in W \mid \exists S \in \mathcal{O}(v), ([\![A]\!] \bullet S \wedge \varnothing \not\bullet S) \}$

(by definition of $[\![\bot]\!]$, definition 2.3.3(item 2))

iff $w \in \{ v \in W \mid \exists S \in \mathcal{O}(v), ([\![A]\!] \bullet S) \}$

iff $\exists S \in \mathcal{O}(w), ([\![A]\!] \bullet S)$

iff $[\![A]\!] \bullet \bigcup \mathcal{O}(w)$

iff $w \in \left\{ v \in W \;\middle|\; [\![A]\!] \bullet \bigcup \mathcal{O}(v) \right\}$.

12.

$w \in [\![\Diamond\!\!\cdot\, A]\!]$

iff $w \in [\![A \preccurlyeq \top]\!]$

(by definition of $\Diamond\!\!\cdot\, A$, definition 2.1.6)

iff $w \in \{\, v \in W \mid \forall S \in \mathcal{O}(v), ([\![\top]\!] \bullet S \to [\![A]\!] \bullet S) \,\}$

(by definition of $[\![A \preccurlyeq B]\!]$, definition 2.3.3(item 4))

iff $\forall S \in \mathcal{O}(w), ([\![\top]\!] \bullet S \to [\![A]\!] \bullet S)$

iff $\forall S \in \mathcal{O}(w), (((W \cap S) \neq \varnothing) \to [\![A]\!] \bullet S)$

(by definition of $\bullet$, definition 2.3.1)

iff $\forall S \in \mathcal{O}(w), (S \neq \varnothing \to [\![A]\!] \bullet S)$

iff $w \in \{\, v \in W \mid \forall S \in \mathcal{O}(v), (S \neq \varnothing \to [\![A]\!] \bullet S) \,\}$.

13.

$w \in [\![\boxdot\, A]\!]$

iff $w \in [\![\top \prec \neg A]\!]$

(by definition of $\boxdot\, A$, definition 2.1.6)

iff $w \in \{\, v \in W \mid \exists S \in \mathcal{O}(v), ([\![\top]\!] \bullet S \wedge [\![\neg A]\!] \not\bullet S) \,\}$

(by theorem of $[\![A \prec B]\!]$, theorem 2.3.5(item 8))

iff $\exists S \in \mathcal{O}(w), ([\![\top]\!] \bullet S \wedge [\![\neg A]\!] \not\bullet S)$

iff $\exists S \in \mathcal{O}(w), ((W \cap S \neq \varnothing) \wedge (S \subseteq [\![A]\!]))$

(by lemma 8.1.1 and by definition of $\bullet$, definition 2.3.1)

iff $\exists S \in \mathcal{O}(w), ((S \neq \varnothing) \wedge (S \subseteq [\![A]\!]))$

iff $w \in \{\, v \in W \mid \exists S \in \mathcal{O}(v), (S \neq \varnothing) \wedge (S \subseteq [\![A]\!]) \,\}$.

14.

$w \in [\![A \mathrel{\Box\!\!\Rightarrow} B]\!]$

iff $w \in [\![((A \wedge B) \prec (A \wedge \neg B))]\!]$

(by definition of $A \mathrel{\Box\!\!\Rightarrow} B$, definition 2.1.6)

iff $w \in \{\, v \in W \mid \exists S \in \mathcal{O}(v), ([\![A \wedge B]\!] \bullet S \wedge [\![A \wedge \neg B]\!] \not\bullet S) \,\}$

(by theorem of $[\![A \prec B]\!]$, theorem 2.3.5(item 8))

iff $\exists S \in \mathcal{O}(w), ([\![A \wedge B]\!] \bullet S \wedge [\![A \wedge \neg B]\!] \not\bullet S)$

iff $\exists S \in \mathcal{O}(w), (([\![A]\!] \cap [\![B]\!]) \bullet S \wedge ([\![A]\!] \cap [\![\neg B]\!]) \not\bullet S)$

iff $\exists S \in \mathcal{O}(w),$

$(([\![A]\!] \cap [\![B]\!] \cap S \neq \varnothing) \wedge ([\![A]\!] \cap [\![\neg B]\!] \cap S = \varnothing))$

(by definition of $\not\bullet$, definition 2.3.1)

iff $\exists S \in \mathcal{O}(w), (([\![A]\!] \cap [\![B]\!] \cap S \neq \varnothing) \wedge [\![\neg B]\!] \not\bullet ([\![A]\!] \cap S))$

iff $\exists S \in \mathcal{O}(w), (([\![A]\!] \cap [\![B]\!] \cap S \neq \varnothing) \wedge (([\![A]\!] \cap S) \subseteq [\![B]\!]))$

(by lemma 8.1.1)

iff $\exists S \in \mathcal{O}(w), (([\![A]\!] \cap S \neq \varnothing) \wedge (([\![A]\!] \cap S) \subseteq [\![B]\!]))$

iff $\exists S \in \mathcal{O}(w), (([\![A]\!] \bullet S) \wedge (([\![A]\!] \cap S) \subseteq [\![B]\!]))$

iff $w \in \left\{ v \in W \;\middle|\; \begin{array}{l} \exists S \in \mathcal{O}(v), \\ \quad (([\![A]\!] \bullet S) \wedge (([\![A]\!] \cap S) \subseteq [\![B]\!])) \end{array} \right\}.$

15. First we will need a lemma.

Lemma 8.1.3.
Given a model with a generalized valuation, sentences A and B, and a set $S \subseteq W$, TFAE:

a) $([\![A]\!] \cap [\![\neg B]\!]) \bullet S \to ([\![A]\!] \cap [\![B]\!]) \bullet S.$

b) $[\![A]\!] \bullet S \to ([\![A]\!] \cap [\![B]\!]) \bullet S.$

Proof. First we prove that $(a) \to (b)$. Suppose (a), suppose $[\![A]\!] \bullet S$. Then there exists a $w \in [\![A]\!] \cap S$. Now we have two cases, either $w \in [\![B]\!]$ or $w \in [\![\neg B]\!]$. In the former case we can immediately conclude $([\![A]\!] \cap [\![B]\!]) \bullet$

S and in the latter case we get that $(\llbracket A \rrbracket \cap \llbracket \neg B \rrbracket) \bullet S$ and by applying our assumption, (a), we get $(\llbracket A \rrbracket \cap \llbracket B \rrbracket) \bullet S$. This gives us ($b$) and therefore $(a) \to (b)$.

Next we prove that $(b) \to (a)$. Suppose (b), suppose $(\llbracket A \rrbracket \cap \llbracket \neg B \rrbracket) \bullet S$. Hence we have that there is a $w \in \llbracket A \rrbracket \cap S$. So we can just apply ($b$) to get $(\llbracket A \rrbracket \cap \llbracket B \rrbracket) \bullet S$. This gives us ($a$) and thus we have that $(b) \to (a)$. $\square$

Now we can prove our statement.

$w \in \llbracket A \diamond\!\!\Rightarrow B \rrbracket$

iff $w \in \llbracket ((A \wedge B) \preccurlyeq (A \wedge \neg B)) \rrbracket$

(by definition of $A \diamond\!\!\Rightarrow B$, definition 2.1.6)

iff $w \in \{ v \in W \mid \forall S \in \mathcal{O}(v), (\llbracket A \wedge \neg B \rrbracket \bullet S \to \llbracket A \wedge B \rrbracket \bullet S) \}$

(by definition of $\llbracket A \preccurlyeq B \rrbracket$, definition 2.3.3(item 4))

iff $\forall S \in \mathcal{O}(w), (\llbracket A \wedge \neg B \rrbracket \bullet S \to \llbracket A \wedge B \rrbracket \bullet S)$

iff $\forall S \in \mathcal{O}(w), ((\llbracket A \rrbracket \cap \llbracket \neg B \rrbracket) \bullet S \to (\llbracket A \rrbracket \cap \llbracket B \rrbracket) \bullet S)$

iff $\forall S \in \mathcal{O}(w), (\llbracket A \rrbracket \bullet S \to (\llbracket A \rrbracket \cap \llbracket B \rrbracket) \bullet S)$

(by lemma 8.1.3)

iff $w \in \{ v \in W \mid \forall S \in \mathcal{O}(v), (\llbracket A \rrbracket \bullet S \to (\llbracket A \rrbracket \cap \llbracket B \rrbracket) \bullet S) \}$.

16.

$$w \in [\![A \,\Box\!\!\rightarrow B]\!]$$

iff $w \in [\![(\Diamond A \rightarrow (A \,\Box\!\!\Rightarrow B))]\!]$

(by definition of $A \,\Box\!\!\rightarrow B$, definition 2.1.6)

iff $w \in [\![\neg \Diamond A]\!] \cup [\![A \,\Box\!\!\Rightarrow B]\!]$

(by theorem of $[\![A \rightarrow B]\!]$, theorem 2.3.5(item 3))

iff $w \notin [\![\Diamond A]\!]$

or $w \in [\![A \,\Box\!\!\Rightarrow B]\!]$

iff $w \notin \left\{ v \in W \;\middle|\; [\![A]\!] \bullet \bigcup \mathscr{O}(v) \right\}$

or $w \in \left\{ v \in W \;\middle|\; \begin{array}{l} \exists S \in \mathscr{O}(v), \\ \quad (([\![A]\!] \bullet S) \wedge (([\![A]\!] \cap S) \subseteq [\![B]\!])) \end{array} \right\}$

(by theorem of $[\![\Diamond A]\!]$, theorem 2.3.5(item 10))

(by theorem of $[\![A \,\Box\!\!\Rightarrow B]\!]$, theorem 2.3.5(item 14))

iff not $\left[[\![A]\!] \bullet \bigcup \mathscr{O}(w) \right]$

or $\exists S \in \mathscr{O}(w), (([\![A]\!] \bullet S) \wedge (([\![A]\!] \cap S) \subseteq [\![B]\!]))$

iff $[\![A]\!] \bullet \bigcup \mathscr{O}(w) \rightarrow$

$\exists S \in \mathscr{O}(w), (([\![A]\!] \bullet S) \wedge (([\![A]\!] \cap S) \subseteq [\![B]\!]))$

iff $w \in \left\{ v \in W \;\middle|\; \begin{array}{l} [\![A]\!] \bullet \bigcup \mathscr{O}(v) \rightarrow [\exists S \in \mathscr{O}(v), \\ \quad (([\![A]\!] \bullet S) \wedge (([\![A]\!] \cap S) \subseteq [\![B]\!]))] \end{array} \right\}.$

17.

$w \in [\![A \diamondto B]\!]$

iff $w \in [\![(\Diamond A \wedge (A \diamondRightarrow B))]\!]$

(by definition of $A \diamondto B$, definition 2.1.6)

iff $w \in [\![\Diamond A]\!] \cap [\![A \diamondRightarrow B]\!]$

(by theorem of $[\![A \wedge B]\!]$, theorem 2.3.5(item 5))

iff $w \in [\![\Diamond A]\!]$

and $w \in [\![A \diamondRightarrow B]\!]$

iff $w \in \left\{ v \in W \;\middle|\; [\![A]\!] \bullet \bigcup \mathcal{O}(v) \right\}$

and $w \in \left\{ v \in W \;\middle|\; \begin{array}{l} \forall S \in \mathcal{O}(v), \\ \quad ([\![A]\!] \bullet S \to ([\![A]\!] \cap [\![B]\!]) \bullet S) \end{array} \right\}$

(by theorem of $[\![\Diamond A]\!]$, theorem 2.3.5(item 10))

(by theorem of $[\![A \diamondRightarrow B]\!]$, theorem 2.3.5(item 15))

iff $[\![A]\!] \bullet \bigcup \mathcal{O}(w)$

and $\forall S \in \mathcal{O}(w), ([\![A]\!] \bullet S \to ([\![A]\!] \cap [\![B]\!]) \bullet S)$

iff $[\![A]\!] \bullet \bigcup \mathcal{O}(w) \wedge \forall S \in \mathcal{O}(w), ([\![A]\!] \bullet S \to ([\![A]\!] \cap [\![B]\!]) \bullet S)$

iff $w \in \left\{ v \in W \;\middle|\; \begin{array}{l} [\![A]\!] \bullet \bigcup \mathcal{O}(v) \wedge \forall S \in \mathcal{O}(v), \\ \quad ([\![A]\!] \bullet S \to ([\![A]\!] \cap [\![B]\!]) \bullet S) \end{array} \right\}.$

□

8.2 Proofs About Various Axioms

Here we prove the remaining theorems for theorem 3.4.3.

Theorem 8.2.1.
The following are true:

1. $\vdash_V \Box A \to (\diamond\!\!\cdot\, A \vee \neg(\top \prec \bot))$.

2. $\vdash_V \boxdot A \to \diamond\!\!\cdot\, A$.

3. $\vdash_{VC}$ ***N***.

4. $\vdash_{VN} (\Box A \vee \boxdot A) \rightarrow \Diamond\!\!\cdot\, A$.

5. $\vdash_{VC}$ ***W***.

6. $\vdash_{VW}$ ***T***.

7. $\vdash_{VT}$ ***N***.

8. $\vdash_{VA_1}$ ***4***.

9. $\vdash_{VA_2}$ ***5***.

Proof. We prove each statement separately.

2. We provide a derivation of $\vdash_{\mathbf{V}} \boxdot A \rightarrow \Diamond\!\!\cdot\, A$.

1.	$\vdash_{\mathbf{V}} \boxdot A \rightarrow$ $([(A \preccurlyeq \top) \vee (\neg A \preccurlyeq \top)] \rightarrow$ $[\boxdot A \wedge [(A \preccurlyeq \top) \vee (\neg A \preccurlyeq \top)]])$	Taut
2.	$\vdash_{\mathbf{V}} [\boxdot A \wedge [(A \preccurlyeq \top) \vee (\neg A \preccurlyeq \top)]] \rightarrow$ $[(\boxdot A \wedge (A \preccurlyeq \top)) \vee (\Box A \wedge (\neg A \preccurlyeq \top))]$	Taut
3.	$\top \rightarrow (A \vee \neg A)$	Taut
4.	$[(A \preccurlyeq \top) \vee (\neg A \preccurlyeq \top)]$	Cp,3
5.	$\vdash_{\mathbf{V}} \boxdot A \rightarrow$ $[(\boxdot A \wedge (A \preccurlyeq \top)) \vee (\boxdot A \wedge (\neg A \preccurlyeq \top))]$	Pl,1,2,4
6.	$\vdash_{\mathbf{V}} \boxdot A \rightarrow$ $[(\boxdot A \wedge (A \preccurlyeq \top)) \vee ((\top \prec \neg A) \wedge (\neg A \preccurlyeq \top))]$	Defn of $\boxdot$
7.	$\vdash_{\mathbf{V}} ((\top \prec \neg A) \wedge (\neg A \preccurlyeq \top)) \rightarrow (\top \prec \top)$	By 3.3.2(5)
8.	$\vdash_{\mathbf{V}} \neg(\top \prec \top)$	By 3.3.2(2)
9.	$\vdash_{\mathbf{V}} \boxdot A \rightarrow (A \preccurlyeq \top)$	Pl,6,7,8
10.	$\vdash_{\mathbf{V}} \boxdot A \rightarrow \Diamond\!\!\cdot\, A$	Defn of $\Diamond\!\!\cdot$

Classical tautologies used:

$$p \to [q \to (p \wedge q)]$$

$$[p \wedge (q \vee r)] \to [(p \wedge q) \vee (p \wedge r)]$$

$$\top \to (p \vee \neg p).$$

3. We provide a derivation of $\vdash_{\mathbf{VC}} \mathbf{N}$.

1.	$\vdash_{\mathbf{VC}} \Diamonddot \bot \to \bot$	**C**
2.	$\vdash_{\mathbf{VC}} (\bot \preccurlyeq \top) \to \bot$	Defn of $\Diamonddot$
3.	$\vdash_{\mathbf{VC}} \top \prec \bot$	Defn of $\prec$

Since every instance of **N** is of the form $\top \prec \bot$, then $\vdash_{\mathbf{VC}} \mathbf{N}$.

4. We provide a derivation of $\vdash_{\mathbf{VN}} (\Box A \vee \boxdot A) \to \Diamonddot A$.

1.	$\vdash_{\mathbf{VN}} \top \prec \bot$	**N**
2.	$\vdash_{\mathbf{VN}} \Box A \to (\Diamonddot A \vee \neg(\top \prec \bot))$	By 3.4.3(1)
3.	$\vdash_{\mathbf{VN}} \Box A \to \Diamonddot A$	PL,1,2
4.	$\vdash_{\mathbf{VN}} \boxdot A \to \Diamonddot A$	By 3.4.3(2)
5.	$\vdash_{\mathbf{VN}} (\Box A \vee \boxdot A) \to \Diamonddot A$	PL,3,4

5. We provide a derivation of $\vdash_{\mathbf{VC}} \mathbf{W}$.

1.	$\vdash_{\mathbf{VC}} \Diamonddot A \to A$	**C**
2.	$\vdash_{\mathbf{VC}} (\Box A \vee \boxdot A) \to \Diamonddot A$	By 3.4.3(3) and 3.4.3(4)
3.	$\vdash_{\mathbf{VC}} (\Box A \vee \boxdot A) \to A$	PL,1,2

Since every instance of **W** is of the form $(\Box A \vee \boxdot A) \to A$, then $\vdash_{\mathbf{VC}} \mathbf{W}$.

6. We actually already proved this in our example deduction above, example 3.1.5.

8. We provide a derivation of $\vdash_{\mathbf{VA}_1} \mathbf{4}$.

1.	$\vdash_{\mathbf{VA}_1} (\bot \preccurlyeq \neg A) \to \Box(\bot \preccurlyeq \neg A)$	$\mathbf{A}_1$
2.	$\vdash_{\mathbf{VA}_1} \Box A \to \Box\Box A$	Defn of $\Box$

Since every instance of **4** is of the form $\Box A \rightarrow \Box\Box A$, then $\vdash_{\mathbf{VA}_1}$ **4**.

9. We provide a derivation of $\vdash_{\mathbf{VA}_2}$ **5**.

1.	$\vdash_{\mathbf{VA}_2} (A \prec \bot) \rightarrow \Box(A \prec \bot)$	$\mathbf{A}_2$
2.	$\vdash_{\mathbf{VA}_2} \Diamond A \rightarrow \Box\Diamond A$	Defn of $\Diamond$

Since every instance of **5** is of the form $\Diamond A \rightarrow \Box\Diamond A$, then $\vdash_{\mathbf{VA}_2}$ **5**.

$\Box$

Chapter 9

Conclusion and Future Work

Throughout this thesis we have presented a modern reformulation of Lewis' counterfactual logic. We have provided formal definitions and explicit proofs where they were absent and added additional results where we believed they could add clarity. Some things we introduced that weren't present in Lewis (2001) included:

1. The notion of frames, with several definitions and theorems being reformulated in terms of frames.

2. A simplification to our language down to three connectives (greatly reducing the length of induction proofs).

3. A variety of logical rules and derivations to make the syntax easier to work with.

Soundness and completeness theorems were proven, with an additional proof of strong local soundness. Canonical models were fully presented in detail together with a complete presentation of Krabbe's counterexample.

This formulation is fully general and does not assume any axioms (except where noted explicitly). Anyone who wants to study or develop a particular counterfactual logic using Lewis' approach can do so on top of the machinery presented in this thesis. Additionally, by having everything explicitly presented it should help readers clear up confusion about how the theory works.

The outer and inner modalities do behave similarly to traditional modal logics (though, the inner modality requires the axiom **N** in order to have an inner necessitation rule theorem 3.4.4 and an internal version of Kripke's axiom theorem 3.4.6). More specifically, given a logic, Lewis (2001, Section 6.3 p. 137) defines:

1. the derived outer modal logic as the sublogic obtained by discarding all connectives aside from the outer modalities,[1] and

2. the derived inner modal logic as the sublogic obtained by discarding all connectives aside from the inner modalities.

He then argues that for each axiomatic system he introduced, the derived outer and inner modal logics are equivalent to a corresponding modal logic. In order to formally demonstrate this, one first needs to introduce the theory of modal logic, together with a description of several modal logics, and then they need to show that any theorem of an inner or outer modal logic is a theorem of the corresponding modal logic and vice versa.

Unfortunately it would not be feasible to give a presentation of these details in this setting. It would be interesting to further develop the relationship between counterfactual logic and modal logic. In some sense, one may think of counterfactual logic as a version of modal logic where the accessibility relation has been replaced with a structure that provides more granular information. This is easy to see by noticing that a traditional modal logic accessibility relation, $R \subseteq (W \times W)$, can be obtained for the outer modality as follows,

$$\forall w, v \in W, (w, v) \in R \text{ iff } v \in \bigcup \mathcal{O}(w).$$

The accessibility relation may be used to determine the behavior of the outer modalities, even though it contains far less information than the system of spheres. However, it is not able to tell us anything about any of our counterfactual or comparative possibility statements. Perhaps one could find cases

[1]Technically speaking, we've defined our underlying language a little differently from Lewis. For us, inner and outer modalities are defined connectives in terms of $\preccurlyeq$. So for us we would discard all instances of $\prec$ that are not equivalent to the modalities we're interested in and we would have to rely on theorems about said modalities in our computations.

where it may be more beneficial to study a particular counterfactual logic in place of a particular modal logic. To this end it would be worthwhile exploring more counterfactual axioms that can be used to derive known modal axioms (not merely importing modal axioms into counterfactual logic as is the case with **T** and others).

Topological models for modal logic were introduced long ago by McKinsey and Tarski (1944) and there has been a wealth of literature on the subject. There are two ways of producing topological models for modal logic, one involves interpreting $\Diamond$ in terms of the topological interior operator and the other involves interpreting $\Box$ as the derived set operator (i.e., the limit points of a set). In each case, the topological models obtained correspond to certain subfamilies of modal logics.

Unfortunately, there has been very little literature in this vein for for Lewis' counterfactual logic. Marti and Pinosio (2014) developed an unusual sort of topological models for a different approach to counterfactual logic using conditional logic. Unfortunately, their approach does not yield one topology over the set of worlds, but rather one topology for each world. It also results in a specialized logic with a lot of properties and modal operators that satisfy a lot of modal axioms. Pacuit (2017) briefly discusses how our sphere model based semantics are an example of neighborhood semantics.

It would be very interesting to have topological semantics and there should be a few ways to obtain them for some counterfactual logics. The simplest way would be to look at axiomatic systems which contain axioms that make either the inner or outer modality behave like the **S4** modal logic. For the outer modality, this would be the axiomatic system **VT4**, and for the inner modality it would be the axiomatic system $\Sigma = \{ \boxdot A \rightarrow A, \boxdot A \rightarrow \boxdot\boxdot A, \mathbf{N} \}$ (note that **N** is necessary to make the inner modality have an inner necessitation rule theorem 3.4.4 and an internal version of Kripke's axiom theorem 3.4.6). From this point, one should be able to obtain a reflexive and transitive relation, which can be used to obtain an Alexandrov topology where $\Box$ can be interpreted as the interior operation. The question then becomes, what do each of these two topologies (one for each pair of modalities) tell us about our counterfactual connectives or comparative possibility? What effects do topological properties have on our topological semantics? One could take this further as well, by searching for stronger axioms written in terms of comparative possibility statements or counterfactual statements that derive the axioms necessary to have topological models. If interesting axioms of the sort exist, then they may provide families of models where one can reason about comparative possibility or counterfactuals in terms of topology.

Another approach, inspired by the obscure notion of a prametric topological space (Arkhangel'skiĭ and Fedorchuk, 2012, Section 2.4 p. 23), would be to assume property **W**, and then define a topology as follows:

1. W is open, and
2. $U \subsetneq W$ is open iff, $\forall w \in U$, $\bigcap(\mathcal{O}(w) - \{ \varnothing \}) \subseteq U$.

Would adding property,

$$\forall w \in W, \forall v \in \mathcal{O}(w), \bigcap(\mathcal{O}(v) - \{ \varnothing \}) \subseteq \bigcap(\mathcal{O}(w) - \{ \varnothing \}),$$

be sufficient for making the $\boxdot$ behave like the interior operator with respect to this topology?

Many other questions arise when considering the notion of topological models for counterfactual logic. For instance, do open sets tell us anything meaningful with respect to similarity between worlds? What are some natural language ways that we can interpret the modalities, and possibly counterfactuals and comparative possibility statements. For instance, could we interpret $\Box A$ as "it will always be true that A," and if so what would that say about our logic and what sort of counterfactual or comparative possibility statements could we make?

Bibliography

Arkhangel'skiĭ, Alexander Vladimirovich and Vitaly Vitalievich Fedorchuk, 2012. *General Topology I: Basic Concepts and Constructions. Dimension Theory*. Berlin: Springer.

Chellas, Brian F., 1975. Basic conditional logic. *Journal of Philosophical Logic* 4(2): 133–153. doi:10.1007/BF00693270.

Dasgupta, Abhijit, 2014. *Set Theory: With an Introduction to Real Point Sets*. New York: Springer.

Krabbe, Erik C. W., 1978. Note on a completeness theorem in the theory of counterfactuals. *Journal of Philosophical Logic* 7(1): 91–93. doi:10.1007/bf00245922.

Lewis, David K., 1971. Completeness and decidability of three logics of counterfactual conditionals. *Theoria* 37(1): 74–85. doi:10.1111/j.1755-2567.1971.tb00061.x.

Lewis, David K., 1973. *Counterfactuals*. Malden, MA: Blackwell.

Lewis, David K., 1981a. Counterfactuals and comparative possibility. In *IFS: Conditionals, Belief, Decision, Chance and Time*, eds. William L. Harper, Robert Stalnaker, and Glenn Pearce, 57–85. Dordrecht: Kluwer. doi:10.1007/978-94-009-9117-0_3.

Lewis, David K., 1981b. Ordering semantics and premise semantics for counterfactuals. *Journal of Philosophical Logic* 10(2): 217–234. doi:10.1007/bf00248850.

Lewis, David K., 2001. *Counterfactuals*. Malden, MA: Blackwell, revised ed.

Marti, Johannes and Riccardo Pinosio, 2014. Topological semantics for conditionals. 115–128. London: College Publications.

McKinsey, J. C. C. and Alfred Tarski, 1944. The algebra of topology. *Annals of Mathematics* 45(1): 141–191. URL http://www.jstor.org/stable/1969080.

Molnar, Christoph, 2020. *Interpretable Machine Learning*. Lulu. com.

Nute, Donald, 1980. *Topics in Conditional Logic*. Dordrecht: Springer Netherlands. doi:10.1007/978-94-009-8966-5.

Pacuit, Eric, 2017. *Neighborhood Semantics for Modal Logic*. Cham: Springer.

Pearl, Judea and Dana Mackenzie, 2018. *The Book of Why: The New Science of Cause and Effect*. New York: Basic Books.

Ribeiro, Marco Tulio, Sameer Singh, and Carlos Guestrin, 2018. Anchors: High-precision model-agnostic explanations. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence*, 1527–1535. Association for the Advancement of Artificial Intelligence. doi:10.1609/aaai.v32i1.11491.

Segerberg, Krister, 1989. Notes on conditional logic. *Studia Logica* 48(2): 157–168. doi:10.1007/BF02770509.

Sider, Theodore, 2010. *Logic for Philosophy*. Oxford: Oxford University Press.

Stalnaker, Robert C., 1968. A theory of conditionals. In *Studies in Logical Theory*, ed. Nicholas Rescher, 98–112. Oxford: Blackwell.

Stalnaker, Robert C. and Richmond H. Thomason, 1970. A semantic analysis of conditional logic. *Theoria* 36(1): 23–42. doi:10.1111/j.1755-2567.1970.tb00408.x.

Unterhuber, Matthias, 2013. *Possible Worlds Semantics for Indicative and Counterfactual Conditionals? A Formal Philosophical Inquiry into Chellas–Segerberg Semantics*. Frankfurt: Ontos. doi:10.1515/9783110323665.

Weiss, Yale, 2019. Frontiers of conditional logic. Ph.D. thesis, The Graduate Center, City University of New York, New York. URL https://academicworks.cuny.edu/gc_etds/2964.

Zach, Richard, 2019. *Boxes and Diamonds: An Open Introduction to Modal Logic*. Calgary: Open Logic Project. URL https://bd.openlogicproject.org/.