Introducing a Logic of *False Lemmas*: Towards an Epistemic Logic of Grounded Belief

Matthew P. Wampler-Doty

Abstract

The purpose of this essay is to introduce a naïve concept of justification from the epistemology literature into epistemic logic. To this end, I first derive philosophical principles regarding the epistemology of *idealized agents* commonly employed in the epistemic logic literature. I arrive at criteria for knowledge similar to the solution proposed by Clark (1963) to the Gettier problem, known as the *no false lemmas* solution. Having provided these preliminary philosophical results, I develop the semantics for a philosophic logic FL, which I refer to as *false lemma* logic in deference to Clark. My approach differs significantly from the approaches present in the existing literature, as instead of considering the sum of an agent's beliefs I consider the *basis* upon which an agent generates their beliefs. I then go on to demonstrate that the *negative doxastic introspection* axiom from epistemic logic is incompatible other possible propositions in FL logic.

Contents

1 Introduction

The purpose of this essay is to introduce a naïve concept of justification from the epistemology literature into epistemic logic. To this end, I first derive philosophical principles regarding the epistemology of *idealized agents* commonly employed in the epistemic logic literature. I arrive at criteria for knowledge similar to the solution proposed by Clark (1963) to the Gettier problem, known as the *no false lemmas* solution. Having provided these preliminary philosophical results, I develop the semantics for a philosophic logic FL, which I refer to as *false lemma* logic in deference to Clark. My approach differs significantly from the approaches present in the existing literature, as instead of considering the sum of an agent's beliefs I consider the *basis* upon which an agent generates their beliefs. I then go on to demonstrate that the *negative doxastic introspection* axiom from epistemic logic is incompatible other possible propositions in FL logic.

1.1 A Note on Philosophical Methodology

This paper assumes the reader has a minimal knowledge of epistemology.

In the course of this essay I will present *definitions*, *propositions*, *theorems*, *lemmas* and *corollaries*. I will present *propositions* when I am primarily interested in philosophy, and *theorems*, *lemmas* and *corollaries* when I am primarily interested in formal logic. I will present *definitions* in both cases.

While I will attempt to present motivation for *definitions*, their content will consist of certain concepts which I will take to be essentially primitive. I do not expect my definitions to be beyond argument. However, for anything I give as a definition, I will be essentially disinterested in arguing about its meaning. I will also present *propositions* which I do not take to be definitive but will none the less be positions I think can be defended, perhaps only weakly.

I take certain concepts to be philosophically primitive, namely the concept of an *agent* and whether a statement is True or False, dependent on circumstances. I assume the reader is already acquainted with these concepts.

Finally, I feel the need to remark on my language. I will use the terms "ideal" and "idealized" interchangeably. I will talk about a belief *basis* and a belief *base* occasionally. It is regrettable that the plural of both of these words is *bases*. I have decided to purposely misspell the plural of basis as *basies*, reflecting pronunciation. I hope the reader will not find this decision too jarring.

1.2 A Note on Logical Notation and Methodology

This essay assumes the reader is acquainted with propositional logic, naïve set theory, formal logic, multi-modal logic and the semantics of modal logic.

I will make use of notation not familiar to all readers. To this end, I provide the following definitions:

Definition 1.2.1
Grammar:

(a) $\neg$ denotes negation

(b) $\rightarrow$ denotes material implication

(c) $\leftrightarrow$ denotes biconditional implication

(d) $\vee$ denotes "or"

(e) $\wedge$ denotes "and"

(f) $\Box$ denotes a modal necessity operator; I will distinguish other modal necessity operators (I will also use $\Box$ to denote the end of a proof)

(g) $\Diamond$ denotes a modal possibility operator; I will distinguish other modal possibility operators

(h) $\overline{S}$ denotes the deductive closure of S

Meta-grammar and Semantics:

(i) $\mathscr{L}_{\mathsf{L}}$ refers to the language of a logic L; where I assume logic L to be predefined

(j) $\mathbb{M}$ refers to a *model* in logic L

(k) True refers to a truth value of "true"

(l) False refers to a truth value of "false"

(m) $\models$ stands for "models" or "semantically entails"

(n) A *validity* will refer to a statement which is always true for a logic L

(o) A *local validity* is a statement which is always true with respect to model $\mathbb{M}$ in logic L

(p) ϕ, ψ, χ, μ denote a formulae in $\mathscr{L}_{\mathsf{L}}$

Set Theory:

(q) Letters in the script font (i.e. $\mathscr{A}, \mathscr{B}, \mathscr{C}, \ldots$) and the fraktur font (i.e. $\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \ldots$) denote sets of formulae in $\mathscr{L}_{\mathsf{L}}$, with the exception of $\mathscr{P}$, as listed below

(r) $\subseteq$ denotes improper set containment (this symbol may be reversed to $\supseteq$ with obvious interpretation)

(s) $A_1 \subseteq A_2 \subseteq \ldots \subseteq A_n$ is referred to as a chain of inclusion

(t) $\cap$ denotes set intersection

(u) $\cup$ denotes set union

(v) $=$ denotes either numerical or set theoretic equality, depending on context

(w) $\in$ denotes set membership

(x) $\mathscr{P}$ denotes the power set operation

(y) $\varnothing$ denotes the empty set

(z) $|S|$ denotes the cardinality of a set S

Other Notation:

(aa) $\iff$ denotes "if and only if" in English

(bb) A strike through a symbol, such as $\neq$ or $\notin$, will denote "not equal" or "not in," and so on

(cc) $\leq$ denotes the numerical "less than or equals to" ordering relation

(dd) $<$ denotes the numerical "less than" strict ordering relation

(ee) $\geq$ denotes the numerical "greater than or equals to" ordering relation

(ff) $>$ denotes the numerical "greater than" strict ordering relation

I will provide further technical definitions of semantic symbols in §4.3.

It might be useful from time to time to talk about the set of local validities in a model $\mathbb{M}$ in logic L; to this end I will present the following definition:

DEFINITION 1.2.2 Let $\mathscr{T}_{\mathbb{M}}$ be the set of local validities in $\mathbb{M}$, that is $\mathscr{T}_{\mathbb{M}} = \{\phi \mid \mathbb{M} \models \phi\}$

For the purposes of this discussion I will restrict myself to formal languages L such that propositional tautologies are validities in any model $\mathbb{M}$. I furthermore assume that it makes sense when discussing logic L to refer to states of affairs in which certain propositions will sometimes be True and sometimes be False. When referring to a certain state of affairs under investigation, I will use the phrases "the current situation" and the "relevant state of affairs" interchangeably.

2 Philosophical Principles

This section is intended to lay down the philosophical foundations that will motivate the epistemic logic I will develop. I define the concepts of an *ideal knower*, an *ideal believer* and an *ideal agent*, the subjects of study in traditional epistemic logic, in terms of mathematical objects I will refer to as *ideal knowledge bases* and *ideal belief bases* respectively. Having provided these definitions, I will go on to philosophically motivate set theoretic bounds for characterizing idealized knowledge in terms of idealized belief. In the process of doing so I will demonstrate that an ideal belief base supervenes on a substructure I call a *belief basis* which I consider to be more philosophically primitive than an ideal belief base. I will define the concept of *possibly grounded belief*, due to Clark (1963), in terms of an ideal agent's belief basis. I will also define what it means to have *fully grounded belief*. Finally, I will present a new chain of inclusion for knowledge, the formal analysis of which I will make the goal of my research.

2.1 Ideal Agents

In this section I lay down the basic concept of an *ideal agent*, the subject commonly studied in the *epistemic logic* literature. By an *ideal agent*, I mean an agent that is both an ideal knower and an ideal believer, which are notions I will define. It should be remarked that the philosophical theory of idealized knowledge I develop builds upon these essential ideas and adds an additional level of detail, namely that of the concept of a *belief basis*.

2.1.1 Ideal Knower

In this section I define the notion of an *ideal knower* in terms of a mathematical structure called an *ideal belief base*, and motivate each point of the definition for the reader. This definition is basically extracted from Hintikka (1962), however it should be noted that Lenzen (1978) and Halpern and Moses (1984) provide a similar definitions. An *ideal knower* is an agent that possesses an *idealized knowledge base*, which may be defined as follows:

Definition 2.1.1 An agent is considered to have an *ideal knowledge base* $\mathscr{K}$ of statements in a language $\mathscr{L}_{\mathsf{L}}$ for a logic L in model $\mathbb{M}$ if and only if $\mathscr{K}$ is the smallest set that has the following properties:

(a) If $\phi \in \mathscr{K}$ and $\phi \to \psi \in \mathscr{K}$, then $\psi \in \mathscr{K}$

(b) $\mathscr{T}_{\mathbb{M}} \subseteq \mathscr{K}$

(c) If $\phi \in \mathscr{K}$, then ϕ must be True in the current state of affairs

As a technical note I will say an ideal knower with knowledge base $\mathscr{K}$ is said to *know* a proposition ϕ if and only if $\phi \in \mathscr{K}$.

I now turn to motivating this definition.

The first part of this definition, Def. 2.1.1(a), may be motivated as follows. Def. 2.1.1(a) states ideal agents know the implications of what they already know. If an ideal knower knows ϕ, and knows $\phi \to \psi$, then she knows ψ also. The natural way to say this in the logic literature is to assert that an ideal knowledge is *closed under material implication*. The sort of idealization I am making here is that it cannot be *revealed* that an ideal knower knows something, in the sense that Socrates reveals knowledge of geometrical propositions in the slave child he interrogates[1] (Plato, 1999a). In addition, ideal knowers are thought to reason on one and the same occasion (for sake of simplicity), and hence are thought to be able to perform any logical computation in arbitrarily small amounts of time. I follow Hintikka (1962) in making this simplifying assumption, which forms the foundation of traditional static epistemic logic.

I move on now to motivating Def. 2.1.1(b), the second part of the definition. Def. 2.1.1(b) states that ideal knower is intended to have background knowledge of statements which are always True

[1]Specifically, I mean no philosopher like Socrates could get an agent to admit to not knowing some proposition ϕ, and then reveal through interrogation that ϕ is a direct consequence of other things the ideal knower reports to know. As such, one could say that ideal knowers are immune to ἔλενχος (*elenkhos*), or the Socratic method of teaching.

in the model $\mathbb{M}$ she is reasoning within. For the most part, logicians and computer-scientists like to hold that ideal knowers know at least all local validities (Halpern, 1999; Hintikka, 1962; Lenzen, 1978; Meyer and van der Hoek, 1995). Minimally, if an ideal knower only knew $(\phi \rightarrow (\psi \rightarrow \phi))$, $((\phi \rightarrow (\psi \rightarrow \chi)) \rightarrow ((\phi \rightarrow \psi) \rightarrow (\phi \rightarrow \psi)))$, $((\neg\phi \rightarrow \neg\psi) \rightarrow ((\neg\phi \rightarrow \psi) \rightarrow \phi))$, they would know any propositional tautology by closure under material implication, that is by Def. 2.1.1(a). As I am restricting this discussion to only languages in which propositional tautologies are validities, and hence local validities, propositional tautologies would certainly make up part of any ideal knowledge base. Ideal knowers are of course thought to have knowledge that does not consist of background knowledge alone, however this knowledge has historically been very difficult to characterize.

Finally, I turn to motivating the last part of the definition, Def. 2.1.1(c). Def. 2.1.1(c) states that an ideal agent cannot posses knowledge of falsity. For any situation it seems strange to say that one may know a falsehood in that situation. The principle that in any situation if one knows some proposition ϕ, then ϕ is true goes back to at least Plato (1999a) and possibly Parmenides. However, sometimes one hears certain people talk about *knowledge* and *belief* interchangeably; depending on the circumstances the distinction is not always that important. It may be considered possible to believe a falsehood, and in such a setting that conflates knowledge with belief it may then be possible to "know" a falsehood. In this essay, however, I have decided to make a distinction between knowledge and belief, as is commonly done in epistemic logic and epistemology in general.

2.1.2 Ideal Believer

In this section I present and motivate a definition of *idealized believer* in terms of an *idealized belief base*. This definition again reflects the epistemic logic literature (Halpern and Moses, 1984; Hintikka, 1962; Lenzen, 1978). As was the case for an idealized knower, an idealized believer is an agent that possesses an ideal belief base. I proceed to define an ideal belief base as follows:

Definition 2.1.2 An agent is considered to have an *ideal belief base* $\mathscr{B}$ of propositions in a language $\mathscr{L}$ in a model $\mathbb{M}$ if and only $\mathscr{B}$ is the smallest set that has the following properties:

(a) If $\phi \in \mathscr{B}$ and $\phi \rightarrow \psi \in \mathscr{B}$, then $\psi \in \mathscr{B}$

(b) $\mathscr{T}_{\mathbb{M}} \subseteq \mathscr{B}$

(c) If $\neg\phi \in \mathscr{T}_{\mathbb{M}}$, then $\phi \notin \mathscr{B}$[2]

As was the case for an ideal knower, an ideal believer with belief base $\mathscr{B}$ is said to *believe* a proposition ϕ if an only if $\phi \in \mathscr{B}$.

I again turn to motivating this definition. Defs. 2.1.1 and 2.1.2 are almost identical, except for the truth condition in Def. 2.1.1(c) is dropped and a *sanity* condition is substituted in as Def. 2.1.2(b)(see Meyer and van der Hoek, 1995). Because of this similarity, and to avoid overlap, I will only motivated the sanity condition on ideal believers (Def. 2.1.2(b)).

[2]The symbol $\perp$, called *falsum*, denotes any ϕ such that $\neg\phi \in \mathscr{T}_{\mathbb{M}}$; this statement can be abbreviated $\perp \notin \mathscr{B}$.

The motivation Def. 2.1.2(b) runs as follows. Following Hintikka (1962), an ideal believer holds all and only *defensible* positions. In addition to being immune to having a philosopher reveal in her beliefs she must hold in virtue of other beliefs, an ideal believer is immune to having a philosopher reveal her beliefs to be in contradiction[3]. Just as in the ideal knower case, it is assumed that the ideal believer is reasoning over a small time span of relative stability, and deducing conclusions arbitrarily quickly.

An note should be made about the difference between ideal knowers and ideal believers. It might be thought that it is necessary to have Def. 2.1.2(c), the sanity condition for beliefs, as part of Def. 2.1.1, the definition of knowledge. To see why this is not necessary, note that while an ideal believer can believe a falsehood, albeit contingent falsehoods, an ideal knower never possesses a falsehood or anything that would lead to a falsehood in their knowledge base. Hence, an ideal knower never has a *falsum* in their knowledge base, because *falsums* are naturally False and hence forbidden to be in ideal knowledge bases by Def. 2.1.1(c). so an analogue of Def. 2.1.2(c) for knowledge is unecessary.

2.2 Bounding Knowledge in Terms of Belief

In this section I illustrate how one may consider an ideal agent that is both an ideal knower and an ideal believer to have a knowledge base bounded in terms of subsets of her belief base. By bounded, I mean that one may construct a chain of inclusion of the mathematical structures I have already defined. In §2.2.1 I provide the bounds commonly employed in the epistemic logic literature. To further refine this bound, I introduce the concept of a *belief basis*, a mathematical substructure which generates an ideal believer's belief base. With this concept in place, I will define the notion of what it means for a belief to be *possibly grounded*, *possibly specious*, and *fully grounded*. With these further refinements in place I will propose a novel chain of inclusion for knowledge.

Before I move on to presenting my view, I feel a few remarks should be made regarding the language I employ. In the subsequent discussion, I will occasionally refer to a knowledge base and a belief base without qualifying those statements in terms of a logic L with some model $\mathbb{M}$ of that logic in mind. I will implicitly intend such qualification to be present. Also, from this point onward I will frequently omit the term *ideal* when referring to ideal agents. When I write "agent" I will intend an ideal agent and not an agent simpliciter.

2.2.1 The Traditional Chain of Inclusion

In this section I present the common chain of inclusion studied in the traditional epistemic logic literature, which embodies the view that an ideal agent's knowledge is bounded below by tautologies and above by true belief. To show this I will propose that an ideal agent's knowledge base is contained in their belief base, illustrate how their knowledge base is in the subset of true beliefs, and show how the set of local validities, $\mathscr{T}_{\mathbb{M}}$, is contained in both of these sets.

[3]Just as an ideal knower is immune to ἐλενχος, I naturally consider an ideal believer as immune to απορια (*aporia*). I am curious if whether Parmenides in Plato's dialogue *Parmenides* (Plato, 1999b) is supposed to be depicted as such an individual. I am not altogether certain if this is the case as Parmenides ontology is difficult to discern and it may be that he holds statements like $\perp$ to be True.

First, I propose that an agent's knowledge base is contained in their belief base. To motivate this, first observe that it is commonly thought in the philosophical literature that knowledge implies belief. Using the set theoretic notation I have so far developed, this may be stated as:

Proposition 2.2.1 For any agent with knowledge base $\mathscr{K}$ and belief base $\mathscr{B}$,

$$\mathscr{K} \subseteq \mathscr{B} \tag{2.2.1}$$

... where the intended reading is as follows: if an agent possesses some proposition ϕ in their knowledge base, then they possess ϕ in their belief base. This is a common set theoretic interpretation of entailment, which I feel mirrors the philosophical concept appropriately.

Second, I bring to the reader's attention that knowledge is in the set of "true" beliefs about whatever circumstances are currently under investigation. This follows as, if ϕ is knowledge, then ϕ is True (in the current situation) according to Def. 2.1.1. With this in mind, I give the following definition:

Definition 2.2.1 Let $\beta = \{\phi \in \mathscr{B} \mid \phi \text{ is } \mathsf{True} \text{ in the current situation}\}$

... and because $\beta \subseteq \mathscr{B}$, I elaborate on Eq. (2.2.1) and say:

$$\mathscr{K} \subseteq \beta \subseteq \mathscr{B} \tag{2.2.2}$$

Finally, from Def. 2.1.1(b), observe that $\mathscr{T}_{\mathbb{M}} \subseteq \mathscr{K}$. This is consistent with what has been observed, for it may be noted that validities are true, and for any agent all local validities are defined to be in that agent's belief base by Def. 2.1.2(b). Thus one may elaborate on Eq. (2.2.2) to arrive at the following:

$$\mathscr{T}_{\mathbb{M}} \subseteq \mathscr{K} \subseteq \beta \subseteq \mathscr{B} \tag{2.2.3}$$

This amounts to all that researchers in traditional epistemic logic ordinarily concern themselves with.

2.2.2 Belief Basis

In this section I introduce the concept of a belief basis. A belief basis, which I will commonly denote as $\mathfrak{B}$, is a substructure of an ideal belief base $\mathscr{B}$ which *generates* $\mathscr{B}$, similar to how a basis in linear algebra generates a vector space. I introduce this concept, because I feel there is nothing more that can be concluded about the bounds on knowledge, described in Eq. (2.2.3), using the existing framework I have so far presented. *Thus, I will take a belief basis to be more primitive than a belief base; For different agents could share identical belief bases but have different belief basies* (I will demonstrate this).

I will present my concept by first defining deductive closure, then defining and motivating the concept of a belief basis, and finally arguing that a belief basis is more primitive than a belief base.

First, I define the notion of a *deductive closure* as follows:

DEFINITION 2.2.2 The *deductive closure* of $\mathfrak{B}$, denoted $\overline{\mathfrak{B}}$, is the smallest set such that

(*i*) $\mathfrak{B} \subseteq \overline{\mathfrak{B}}$

(*ii*) $\mathscr{T}_{\mathbb{M}} \subseteq \overline{\mathfrak{B}}$

(*iii*) If $\phi \in \overline{\mathfrak{B}}$ and $\phi \to \psi \in \overline{\mathfrak{B}}$, then $\psi \in \overline{\mathfrak{B}}$

For a motivation of deductive closure, I refer the reader to look at any fine logic textbook (for instance, see Kyburg and Teng, 2001).

Next, having defined deductive closure, I may now define a *belief basis* as follows:

DEFINITION 2.2.3 An ideal agent has a *belief basis* $\mathfrak{B}$ of her ideal belief base $\mathscr{B}$ if and only $\overline{\mathfrak{B}} = \mathscr{B}$

As a technical note, the fraktur font (i.e. $\mathfrak{A}, \mathfrak{B}, \mathfrak{C}$, etc.) will always be used for a belief basis and subsets of a belief basis. The fraktur letter for E, which is rendered $\mathfrak{E}$, will not be used as it is easily confused with with the fraktur letter for C, which is rendered $\mathfrak{C}$.

I turn now to motivating the above definition. Recall that $\mathscr{T}_{\mathbb{M}}$ represents, in a sense, background knowledge that an agent has about a model $\mathbb{M}$, as I stated in motivating Def. 2.1.1(*b*). In contrast to background knowledge, a belief basis corresponds to the portion of an agent's belief base of which they hold as *evidence regarding the current circumstances they are in*, which might or might not be True. Readers familiar with epistemology will be able to identify this assumption as *foundationalist*. I imagine that an ideal agent naturally takes all of her evidence and combines it with all of her background experience and knowledge to deduce all of her beliefs, which represents every possible argument she can make.

Finally, I argue that a belief basis is more primitive than a belief base. As stated above, it may be remarked that different belief basies can generate identical belief bases — $\overline{\{p, p \to q\}} = \overline{\{p, q\}}$ for instance (assuming that p and q are sentence letters in $\mathbb{M}$). A belief base actually generates an equivalence class of generating basies; to see this, construct $[\![\mathscr{B}]\!] = \{\mathfrak{B} \subseteq \mathscr{L}_{\mathsf{L}} \mid \overline{\mathfrak{B}} = \mathscr{B}\}$. From this observation I conclude the following: *studying belief bases abstracts away detail when studying idealized agents*. For the remainder of this essay, belief basies will take the place of belief bases as the when studying ideal agents.

2.2.3 An Additional Bound on Knowledge: Knowledge entails ***Possibly Grounded*** Belief

In this section I provide my first refinement on the traditional chain of inclusion employed in epistemic logic. I first define what it means for an agent to have a *possibly grounded* belief in terms of subsets of her belief basis $\mathfrak{B}$, borrowing a concept originally due to Clark (1963). I go on to suggest a philosophical principle about possibly grounded beliefs, which I call the *justification principle*. I go on to propose that knowledge entails possibly grounded belief, and motivate this proposition with examples inspired by Gettier (1963). Having established this principle, I will provide a refinement on the traditional chain of inclusion for knowledge.

As a remark, while I draw upon Gettier (1963) for inspiration, I do not intend to solve the so called *Gettier problem.* I follow Zagzebski (1994) in holding that the Gettier problem has no solution. I will ignore it completely in this essay.

I turn now to defining what is meant for an agent to hold a possibly grounded belief. An agent is said to have a possibly grounded belief in ϕ if and only if ϕ follows from true beliefs in the agent's belief basis and background knowledge, embodied by $\mathscr{T}_{\mathbb{M}}$. To this end, I give the following definitions:

DEFINITION 2.2.4 For any agent with belief basis $\mathfrak{B}$:

(a) A set $\mathfrak{C}$ is said to be *grounded* if and only if for every $\psi \in \mathscr{L}_{\mathsf{L}}$, if $\psi \in \mathfrak{C}$, then ψ is True in the relevant circumstances

(b) An agent is said to have *belief in ϕ which is possibly grounded* if and only if there is some $\mathfrak{C} \subseteq \mathfrak{B}$ which is grounded and $\phi \in \overline{\mathfrak{C}}$.

I now turn to motivating this definition.

I will start by providing motivation for Def. 2.2.4(a). For any agent with belief basis $\mathfrak{B}$, if $\mathfrak{C} \subseteq \mathfrak{B}$ and $\mathfrak{C}$ is grounded, then any argument from beliefs in $\mathfrak{C}$ (and background knowledge) the agent can make is a *sound* argument. A *sound* argument, the reader will recall, is a valid argument that possesses only premises that are True in the relevant circumstances. In fact, for every statement ϕ in $\overline{\mathfrak{C}}$, the agent is able to make a *sound* argument for ϕ. Also, it should be noted that if $\mathfrak{C} \subseteq \mathfrak{B}$, then $\overline{\mathfrak{C}} \subseteq \overline{\mathfrak{B}}$. Thus if $\mathfrak{C} \subseteq \mathfrak{B}$ is grounded, and $\phi \in \overline{\mathfrak{C}}$, then $\phi \in \overline{\mathfrak{B}}$; which is consistent with the idea that if an agent can make an argument from a subset of her belief basis, then she believes it.

Next, I turn to motivating Def. 2.2.4(b). The definition amounts to saying that the agent can make at least one sound argument for ϕ from what she believes. It should be noted that she might be able to make more than one sound argument from her belief basis; however I will return to this subject in §2.2.4.

Having provided a definition for possibly grounded belief, I turn now to what I call the *justification principle*:

PROPOSITION 2.2.2 (*The Justification Principle*) If an agent has a possibly grounded belief in ϕ, then ϕ is True in the current circumstances

...this principle amounts to saying that if an agent can make a sound argument from her beliefs to some conclusion, then that conclusion is Truein the current circumstances. I invite the reader to observe that many philosophers hold this principle in some form or another. Furthermore, if a statement is False in the current circumstances, then it is neither (A) knowledge (by 2.1.1(c)) nor (B) possibly grounded (by Prop. 2.2.2, *the justification principle*).

I now move on to presenting the philosophical principle that knowledge entails *possibly grounded* belief. This is stated in the following proposition:

PROPOSITION 2.2.3 If an agent knows ϕ, then the agent has a possibly grounded belief in ϕ

To motivate this proposition, I present the following thought experiment, inspired by Gettier (1963). Let Beth be an agent, and suppose she has some friends Cayley and Dara. Beth is ignorent of what Dara is doing this afternoon, so Beth asks Cayley, who is ordinarily a reliable source of information. Cayley tells Beth "If Dara is downtown, then she is at a café" and "Dara is downtown." However, suppose Dara is uptown and at a café; and Cayley happened to uncharacteristically fail to give Beth accurate information regarding Dara's whereabouts.

I will cast this thought experiment into the formalism I have so far developed, in order to illustrate Prop. 2.2.3. For the sake of simplicity I will assume that the statements that Cayley told Beth constitute a belief basis. With this in mind I will denote Beth's belief basis as $\mathfrak{B} = \{$"If Dara is downtown, then she is at a café", "Dara is downtown"$\}$. Note that "Dara is at a café" is in $\overline{\mathfrak{B}}$. It is doubtful that Beth has *knowledge* of the proposition "Dara is at a café"; any argument she can make for why (and she can only make one) relies on something false in her belief basis.

Cayley could give Beth other reasons to conclude Dara is at a café however. Suppose that Cayley says, in addition to what she has told Beth: "If Dara is drinking coffee right now, she is at a café", "Dara is drinking coffee right now", "If Dara took the bus, then she is at a café", "Dara took the bus", and possibly others. However, if it is the case that Dara is drinking tea instead of coffee, and rode her bike instead of taking the bus, and so on, then Best still does not have knowledge of the proposition that "Dara is at a café." Prop. 2.2.3 reflects the intuition about the case for Beth: *no matter how many unsound arguments Beth can make for Dara's presence at a café, if she cannot make a sound argument, then she does not have knowledge, even if she has true belief.*

Taking this thought experiment as inspiration I now present a further refinement to the traditional chain of inclusion: knowledge is possibly grounded belief. Note that all validities are possibly grounded, as $\varnothing \subseteq \mathfrak{B}$, and $\mathscr{T}_{\mathbb{M}} \in \overline{\varnothing}$ by Def. 2.2.2(*ii*). Furthermore, as noted above, propositions that are possibly grounded are beliefs. By the justification principle, they are True. However, recall from the thought experiment that it is not necessarily the case that a true belief is possibly grounded. In the interest of combining these observations I provide the following definition:

Definition 2.2.5 For an agent with a belief basis $\mathfrak{B}$, let $\mathsf{P} = \{\phi \in \overline{\mathfrak{B}} \mid \phi$ is possibly grounded in $\mathfrak{B}\}$

...with this definition, I refine Eq. (2.2.3) and state where the new set P falls in the hierarchy:

$$\mathscr{T}_{\mathbb{M}} \subseteq \mathscr{K} \subseteq \mathsf{P} \subseteq \beta \subseteq \mathscr{B} \tag{2.2.4}$$

2.2.4 Another Bound on Knowledge: *Fully Grounded* Belief entails Knowledge

In this section, I present my final refinement on the chain of inclusion for knowledge. I hold that belief that is *fully grounded* is knowledge. Fully grounded belief corresponds to those beliefs for which the agent can *only* employ sound reasoning to arrive at, in a sense. Fully grounded beliefs have the property that nothing in an agent's experience that is false may serve as justification for such statements.

I will start by first defining what it means for a subset of an agent's belief basis to be *specious*. I will then give a definition for *fully grounded* belief in terms of grounded and specious subsets. I endeavor to carefully motivate this definition, as it may be challenging. I will provide examples of two agents: one that has fully grounded belief in a proposition despite holding a falsehood, and one that has a possibly grounded belief but not a fully grounded belief. I will then give a sufficient criterion for knowledge in terms of fully grounded belief. Finally, I remark on my last revision to the chain of inclusion for knowledge.

I start by providing a definition for what it means for a subset of one's beliefs to be *specious* as follows:

Definition 2.2.6 For any agent with belief basis $\mathfrak{B}$:

(a) $\mathfrak{C}$ is said to be *specious* if and only if there is some $\psi \in \mathfrak{C}$ such that ψ is False in the relevant circumstances

(b) An agent is said to have *belief in ϕ which is possibly specious* if and only if there is some $\mathfrak{C} \subseteq \mathfrak{B}$ which is specious and $\phi \in \overline{\mathfrak{C}}$

I now present motivation for this definition. Specious subsets of an agent's belief basis represent sets of experiential or evidential beliefs that can possibly lead to an unsound argument. Note that any sound argument for a proposition ϕ can be made into an unsound argument with the simple addition of a spurious false premise. Thus, saying that ϕ is possibly specious says very little - ϕ could be true, or it could be false. There is no analogue of the justification principle for specious subsets.

With the concept of a specious subset of a belief basis now provided, I move on to stating what it means for a belief to be *fully grounded*:

Definition 2.2.7 An agent with belief basis $\mathfrak{B}$ is said to have a *fully grounded* belief in ϕ if and only if

(a) ϕ is possibly grounded

(b) For every specious $\mathfrak{C} \subseteq \mathfrak{B}$ such that $\phi \in \overline{\mathfrak{C}}$, there is a grounded subset $\mathfrak{D} \subseteq \mathfrak{C}$ such that $\phi \in \overline{\mathfrak{D}}$

I move now to motivating this definition — I admit it is difficult to understand at first.

I will start by presenting Def. 2.2.7(a). This criterion just says that the agent has a sound argument for ϕ. It is necessary to restrict what is considered a fully grounded belief when the second criterion is satisfied vacuously. In light of Def. 2.2.7(b), Def. 2.2.7(a) could be substituted with the statement "the agent believes ϕ" or "ϕ is possibly specious" without loss of meaning.

Next, I present Def. 2.2.7(b). Assuming that the agent can make an argument for ϕ, this statement says that any unsound argument the agent can make for ϕ can be "fixed," with spurious false premises removed so that a sound argument for ϕ remains. In a sense, if the agent believes ϕ, then she never needs to depend on an unsound premise in any argument she can make for ϕ.

In order to understand the concept of fully grounded belief I will present some examples. I return to Beth, Cayley and Dara from my previous thought experiment. Suppose that later in the evening Beth asks Cayley what Dara is doing. As the reader may recall, I stipulated that Cayley is ordinarily reliable at reporting Dara's activities. Cayley tells Beth the following series of statements

(a) "If Dara is at the library, then she is working on calculus"

(b) "If Dara is working with Eliza, then she is working on calculus"

(c) "If it's monday, then Dara is working on calculus"

(d) "Dara is at the library"

(e) "Dara is working with Eliza"

(f) "It's monday"

(g) "Dara is drinking coffee"

Again, make the simplifying assumption that these statements compose the corpus of the belief basis $\mathfrak{B}$ for Beth. In addition, suppose that all these statements save for (g) are True. Let (g) be false; perhaps Dara is drinking tea that evening. By Def. 2.2.7 Beth has fully grounded belief that "Dara is working on calculus." To see this, take any specious subset of $\mathfrak{B}$ — for instance, $\mathfrak{C} = \{$"Dara is at the library", "Dara is working with Eliza", "If Dara is at the library, then she is working on calculus", "Dara is drinking coffee"$\}$. It is easily seen that $\mathfrak{C}$ is specious, for $\mathfrak{C}$ contains a statement that is false, namely "Dara is drinking coffee". Note that "Dara is working on calculus" $\in \overline{\mathfrak{C}}$. Although Beth could make a possibly unsound argument for Dara's activities, where she might mention "Dara is drinking coffee" spuriously, Beth never *depends* on a false belief to conclude Dara's activities. This is because no argument which involves the statement "Dara is drinking coffee" for why "Dara is working on calculus" that Beth can make will ever be free of a spurious false premise.

I now illustrate how an agent can have a *possibly grounded* belief that is not *fully grounded.* Suppose that Cayley goes on to tell Beth:

(h) "If Dara is drinking coffee, then she is working with Fey"

(i) "If Dara is working with Eliza, then she is working with Fey"

Suppose that Dara is working with Fey. Beth's belief that "Dara is working with Fey" is possibly grounded by not fully grounded. This is because if $\mathfrak{D} = \{$ "Dara is working with Eliza", "If Dara is working with Eliza, then she is working with Fey"$\}$, then $\mathfrak{D}$ is grounded and "Dara is working with Fey"$\in \overline{\mathfrak{D}}$. However, if $\mathfrak{F} = \{$ "Dara is drinking coffee", "If Dara is drinking coffee, then she is working with Fey"$\}$, then "Dara is working with Fey"$\in \overline{\mathfrak{F}}$. It is easily seen that there is no grounded set $\mathfrak{G} \subseteq \mathfrak{F}$ such that "Dara is working with Fey"$\in \overline{\mathfrak{G}}$. One may observe that (sometimes) Beth holds "Dara is working with Fey" as a *false lemma* with no spurious premises, however Beth has other means of arriving at "Dara is working with Fey" which involve sound reasoning and not *false lemmas.*

Having provided and analyzed these examples, I move on to state a sufficient criterion for knowledge for ideal agents. I present the following proposition:

PROPOSITION 2.2.4 If an agent's belief in ϕ is fully grounded, then that agent knows ϕ

If an agent holds a proposition ϕ which is fully grounded, then in a sense any "bad" reasoning the agent may employ for arriving at ϕ has sound reasoning at its core. Despite the fact that the agent might be confused about other propositions, the confusion does not actually affect the soundness of the agent's reasoning regarding ϕ in any way. Thus, if an agent's belief in ϕ is fully grounded, then that agent knows ϕ, for having fully grounded belief in a proposition is a strong condition.

I finally turn to providing a set theoretic bound on knowledge. To this end, I give the following definition:

DEFINITION 2.2.8 For an agent with a belief basis $\mathfrak{B}$, let $\flat = \{\phi \in \overline{\mathfrak{B}} \mid \phi \text{ is fully grounded in } \mathfrak{B}\}$

Provided that $\varnothing$ is a grounded set of any belief basis, then anything that follows from $\varnothing$ is fully grounded, and $\mathscr{T}_{\mathbb{M}} = \overline{\varnothing}$, so the set of all validities is fully grounded. Combining this observation with Eq. (2.2.4), I state my final chain of inclusion:

$$\mathscr{T}_{\mathbb{M}} \subseteq \flat \subseteq \mathscr{K} \subseteq \text{ᚹ} \subseteq \beta \subseteq \mathscr{B} \tag{2.2.5}$$

2.3 Closing Remarks

In this section present a couple of comments in this section regarding the bounds on belief I have provided. I then provide my concluding statement regarding the philosophical analysis I have put forward.

My first remark is on Props. 2.2.3 and 2.2.4. A reader might notice that between these two propositions I have provided necessary and sufficient conditions for knowledge. These propositions leave a lot to be filled in the analysis of knowledge; different researchers in different fields have diverging ideas regarding what knowledge is precisely. To this end I do not intend to give a complete analysis of knowledge in this essay. I leave it to readers interested in application to find definitive criteria for knowledge for ideal agents, provided that they subscribe with what I have to say up to this point.

My second comment is to provide a personal view of how I conceive of ideal agents. I envision an agent's beliefs as being like an scientific subject suitable for dissection. Dissecting an agent's beliefs provides a sense for how that agent thinks. I think this sort of dissection is critical to understanding ideal agents.

Finally, I want to remark that the philosophical analysis I have so far carried out has basically assumed that an epistemic logic framework for agents with basis for belief already exists. I have been unable to find such a reasoning system in the literature; however the literature is quite vast and I do not profess comprehensive familiarity. An epistemic logic that is close to the logic I have in mind is **JTB**, a logic intended for *justified true belief*, due to Voorbraak (1992). Voorbraak's logic is distinguished as it does not permit the "dissection" analysis I am interested in doing with agents. Likewise, Lenzen (1978) has also contributed to logical analysis of justified true belief and belief bases, however his framework also is inadequate for my project. Other deep insights into the nature of knowledge and belief bases are the *knowledge and ignorence* research by Halpern (1997); Halpern and Moses (1984); van der Hoek and Thijsse (1994). Perhaps the closest analysis to my project is due to Fagin et al. (1995), where the authors a view where agents are thought to be in states as well; however the truthfulness of the belief states of agents is not considered. Another approach which is similar in character to the one I have outlined here is due to Artemov and Nogina (2005a,b), who have attempted to introduce a notion of justification to epistemic logic, in order to build up an epistemic logic incorporating *justified true belief*. Their approach essentially is to introduce elements of proof theory to the existing framework; while I think this is a fruitful avenue of research, their approach does not yet allow for analysis for false beliefs. Therefore, I conclude that in order to carry on with my analysis beyond pure philosophy, I will develop a logic of my own to permit the formal reasoning I am interested in. The philosophical observations I have made in this section will be codified in a new epistemic logic I will call FL, the logic of *false lemmas*.

3 Traditional Epistemic Logic

3.1 Introduction

In this section I write about traditional epistemic logic as put forward by Hintikka (1962); Kraus and Lehmann (1986); Lenzen (1978); Meyer and van der Hoek (1995), and others. In this essay I will refer to the simple epistemic logic I am interested in studying as EL. This section is intended to give readers unacquainted with the field a sense of the framework I am elaborating.I will start by presenting the grammar of EL, and then move on to philosophically motivating the basic principles of epistemic logic, before presenting formal semantics. After that, I will present formal semantics, followed by axioms that have been investigated as well as results other researchers have arrived at, including soundness and completeness theorems. I conclude with the problem of *the collapse of belief and knowledge relations.*

This discussion will proceed with only one agent, for the sake of simplicity. It should be remarked that I explicitly imitate Hintikka (1962) in this respect; in deference to Hintikka, my philosophical approach barrows heavily from his original vision of epistemic logic, and less from modern approaches. Despite this, I will endeavor to present my view with modern grammar and semantics, following the presentation in Meyer and van der Hoek (1995), instead of Hintikka's.

3.2 Grammar

In this section I present the language of epistemic logic. Epistemic logic is an extension of propositional logic, just like most other modal logics. That being said, let $\mathbf{P}$ be a set of propositional atoms or constants:

Definition 3.2.1 Let $\mathscr{L}_{\mathsf{EL}}(\mathbf{P})$ be the smallest set such that:

(a) $\mathbf{P} \subseteq \mathscr{L}_{\mathsf{EL}}(\mathbf{P})$

(b) If $\{\phi, \psi\} \subseteq \mathscr{L}_{\mathsf{EL}}(\mathbf{P})$, then $(\phi \wedge \psi) \in \mathscr{L}_{\mathsf{EL}}(\mathbf{P})$

(c) If $\phi \in \mathscr{L}_{\mathsf{EL}}(\mathbf{P})$, then $\{\neg\phi, K\phi, B\phi\} \subseteq \mathscr{L}_{\mathsf{EL}}(\mathbf{P})$

As will be anticipated, the logic that will use this language will be multi-modal.

I fell some notes on notational should be made. In the subsequent discussion, I will omit outside parenthesis. As usual, $\neg(\phi \wedge \neg\psi)$ is abbreviated as $\phi \to \psi$, $\neg(\neg\phi \wedge \neg\psi)$ is abbreviated as $\phi \vee \psi$, and $\phi \leftrightarrow \psi$ as $(\phi \to \psi) \wedge (\psi \to \phi)$. Assume the usual precedence of operations if I omit parenthesis, however I will make an effort never to do so. I also will use the abbreviation $P\phi$ for $\neg K\neg\phi$, which will stand for epistemic possibility, and $C\phi$ for $\neg B\neg\phi$, which stand for doxastic possiblity or conceivability. Finally, I will commonly abbreviate $\mathscr{L}_{\mathsf{EL}}(\mathbf{P})$ to simply $\mathscr{L}_{\mathsf{EL}}$.

3.3 Semantics

In this section I will first present the philosophical motivation for the semantics of epistemic logic. I will then present basic semantics.

3.3.1 Philosophical Motivation for Semantics

In this section I provide philosophical motivation for the semantics of epistemic logic. Epistemic logic is a branch of modal logic; it is similar to many other forms of modal logic in that it employ semantics that involve *possible worlds*[4]. The possible worlds studied in epistemic logic are different from the *elethic* possible worlds which are commonly studied by philosophers interested in metaphysics and modality; accessibility to possible worlds for an agent is considered *subjective*, reflecting that agent's belief or knowledge. This section will consist of an explanation for why an agent might reason about knowledge of the world in terms of possible worlds, as well as a couple of comments on such modeling.

It is necessary to inform the reader that instead of motivating both knowledge and belief, I will only present motivation for belief; the motivation for knowledge is roughly the same.

I now explain why an agent might reason using notions analogous to *possible worlds*. Suppose that Brittany is on a airplane, thinking about what her roommate might be doing at that moment. Perhaps from a discussion with her roommate Brittany had before departure, Brittany holds some propositions regarding her roommate:

(a) "If my roommate is at the dorm, then she is playing on the *Nintendo Wii*"

(b) "If my roommate is at the library, then she is working on game theory homework"

[4]It should be noted that not all modal logics have possible world semantics; for instance, in the proof theory literature there is the modal logic GL developed by Boolos (1993) which rests upon the notion of whether an arithmetical proposition is *provable* as introduced by Gödel (1992) in 1930.

(c) "Either my roommate is at the dorm or at the library, but not both"

The basic idea behind epistemic and doxastic logic is to assume that the above statements regard *possible configurations of the world* that Brittany considers. Brittany holds that the world can be in several possible configurations, called *states*. It is commonly assumed that one of these states is the "true" state, corresponding to reality; this is the state that Brittany actually occupies. I will refer to the true configuration as ω. Beliefs are reflected by a *doxastic accessibility relation*, denoted R_B, similar to elethic modal logic (where the accessibility relation is commonly denoted as R). A state of affairs s' can be thought to be doxastically accessible just in case it is not ruled out by the agent at a state s. Doxastic accessibility is denoted $R_B(s, s')$; epistemic accessibility is similarly denoted $R_K(s, s')$. Statements of belief at a particular state s are then statements about any state s' such that $R_B(s, s')$.

It is instructive to go over an examples to illustrate the above concepts. For instance, (a) states that, in any state s' such that $R_B(\omega, s')$ for Brittany, either Brittany's roommate is at the dorm playing video games or she is not at the dorm at state s'. If I let d denote "Brittany's roommate is at the dorm" and w denote "Brittany's roomate is playing the Nintendo Wii," then one can naturally write (a) as $d \to w$, and the statement "Brittany believes (a)" as $B(d \to w)$. Likewise, (b) states that either (A) Brittany's roommate is at the library doing gametheory homework or (B) she is not at the library, at any doxastically accessible state s'. Just as before, one can codify Brittany's belief in this statement as $B(l \to g)$, with appropriate meanings assigned to l and g respectively. Finally, (c) states that at any doxastically accessible state s' that either it is the case that Brittany's roommate is at the library or at her dorm, but not both. It is natural to think that Brittany will conclude from these beliefs that her roommate is either playing video games or doing game theory[5].

In epistemic logic, statements of belief may address more than just statements about the world; they may also regard propositions and statements about belief or even knowledge. For instance, suppose Brittany held the following proposition:

(d) "My backpack is under the seat in front of me, and I believe it"

Such a statement is considered *introspective*, for it states a belief regarding a belief. The above statement might be denoted $b \wedge B(b)$, and belief in such a statement might be denoted $B(b \wedge B(b))$. This statement says that in any world that is doxastically accessible to Brittany, her backpack is beneath the seat in front of her and she believes it. In a sense, a belief regarding a belief asserts *certitude*; Brittany holds it as a belief regarding her *psychology* that, in any doxastically accessible world, she believes in that world that she her bag is under the seat in front of her. A more radical statement might be:

(e) "Either I believe the movie playing right now is *Gigli* or I believe the movie playing right now is *The Matrix III*"

[5]Note however that in the situation so described, Brittany has no reason to conclude that her roomate is not doing both of these things at the same time; to conclude this will require either background knowledge or further beliefs on the constraints of what her roomate is doing.

Some authors consider whether an agent can believe such statements to be questionable (Halpern, 1997; Halpern and Moses, 1984; Meyer and van der Hoek, 1995; van der Hoek and Thijsse, 1994); I will not discuss this topic here, however.

A comment should be made to fill in necessary gaps I have intentionally left open up to this point in the essay. A *model* $\mathbb{M}$ which I have so far alluded to more or less anonymously in this essay has the intended interpreted of a set of possible worlds in epistemic logic which have doxastic and epistemic accessibility relations connecting them. Local validities are statements which are true in every possible configuration of the system that the agent is modeling; as stated before local validities naturally codify background knowledge. This is because local validities are invariant to all of the configurations the system can take on, including the configuration of the beliefs of the agent. Note that while the discussion in this section has been restricted to discussion of worlds under consideration at a particular world ω, it is assumed that at those worlds the agent has beliefs there as well, which are also governed by the same doxastic accessibility relation. Thus, a model $\mathbb{M}$ consists of a system which is *closed under accessibility relations* (along with other properties).

Finally, I have a remark on the focus of traditional epistemic and doxastic logic. Traditional epistemic logic is restricted to dealing with material implication, conjunction, disjunction, negation, belief and knowledge. Dynamic logic is a modern extension on epistemic logic which addresses subjunctive conditionals and action (see van Ditmarsch et al., 2003). Other exotic extensions exist for probabilistic reasoning (see Kooi, 2003), as well as many other kinds of reasoning (see Meyer and van der Hoek, 1995). I purposefully restrict myself to the static case with one agent in this essay for sake of simplicity; this subject may be elaborated to be as sophisticated as any investigator would like.

3.3.2 Basic Semantics

I submit the following definition of a *Kripke structure*:

Definition 3.3.1 A *Kripke structure* $\mathbb{M}$ is a tuple $\langle \Omega, \pi, R_K, R_B \rangle$ where:

(i) Ω is a non-empty set of *states*

(ii) $\pi : \Omega \to (\mathbf{P} \to \{\mathsf{True}, \mathsf{False}\})$ is a truth assignment to a proposition at a particular world

(iii) $R_K \subseteq \Omega \times \Omega$ and $R_B \subseteq \Omega \times \Omega$

The relation $(\mathbb{M}, s) \models \phi$ (that is, ϕ is satisfied by world s in model $\mathbb{M}$) is defined inductively as follows:

$$\begin{aligned}
(\mathbb{M}, s) \models p &\iff \pi(s)(p) = \mathsf{True} \text{ where } p \in \mathbf{P} \\
(\mathbb{M}, s) \models \phi \wedge \psi &\iff (\mathbb{M}, s) \models \phi \text{ and } (\mathbb{M}, s) \models \psi \\
(\mathbb{M}, s) \models \neg\phi &\iff (\mathbb{M}, s) \not\models \phi \\
(\mathbb{M}, s) \models \Box\phi &\iff (\mathbb{M}, s) \models \phi \text{ for all } s' \text{ with } R_{\Box}(s, s')
\end{aligned}$$

... where $\Box$ is K or B

It is important to talk about what it means for a statement to be valid or satisfied in a Kripke structure and other circumstances. I provide the following definition:

Definition 3.3.2 Let ϕ be an epistemic formula in the language $\mathscr{L}_{\mathsf{FL}}$, and let Φ be a set nonempty set of epistemic formulae in $\mathscr{L}_{\mathsf{FL}}$:

(i) ϕ is *valid in a Kripke model* $\mathbb{M} = \langle \Omega, \pi, R_K, R_B \rangle$, denoted $\mathbb{M} \models \phi$, if for all $s \in \Omega$, $(\mathbb{M}, s) \models \phi$ (such statements are what I call *local validities*)

(ii) ϕ is *valid*, denoted $\models \phi$, if $\mathbb{M} \models \phi$ for all $\mathbb{M}$

(iii) ϕ is *satisfied in* $(\mathbb{M}, s)$ if $(\mathbb{M}, s) \models \phi$

(iv) A set Φ is *valid in a Kripke model* $\mathbb{M}$ if, for all $\phi \in \Phi$, $\mathbb{M} \models \phi$

(v) A set Φ is *valid*, denoted $\models \Phi$ if, for all $\phi \in \Phi$, $\models \phi$

(vi) A set Φ is *satisfied in* $(\mathbb{M}, s)$ if for all $\phi \in \Phi$, $(\mathbb{M}, s) \models \phi$

I will reuse many of these definitions in my discussion of the novel epistemic logic I will develop.

3.4 Characterizing Accessibility Relations

In this section I will characterize the accessibility relations in the semantics of epistemic logic.

3.4.1 The R_K accessibility relation

The only requirement so far made on R_K in the treatment of epistemic logic I have given in this essay is that it satisfy the *truth condition* on knowledge, which is codified in Def. 2.1.1(c). In the epistemic logic literature, the way to ensure this is to make R_K *reflexive*, which is to say that for every world $s \in \Omega$, $R_K(s, s)$. The other parts of the definition of a knowledge base, Defs. 2.1.1(a) and 2.1.1(b), are automatically satisfied; they are features all modal operators posses in ordinary modal logic(see Meyer and van der Hoek, 1995). I will go over this in §3.5.

3.4.2 The R_B accessibility relation

Following the treatment in Kraus and Lehmann (1986); Meyer and van der Hoek (1995); van der Hoek (1991), it is generally thought that $R_B \subseteq R_K$. This ensures that Prop. 2.2.1 is automatically satisfied. It is also assumed that R_B is *serial*, that is to say for every state $s \in \Omega$ that there exists a $s' \in \Omega$ such that $R_B(s, s')$; this ensures that Def. 2.1.2(c) holds. Again, the other features of this definition, namely Defs. 2.1.2(a) and 2.1.2(b), are again achieved automatically.

3.5 Consequences and Validities

In this section I present some consequences and validities of EL. These results are well documented in the epistemic logic literature, so I will not go over their proofs. I will also point out how these results mirror the basic core epistemic logic project I put forward in §2.2.1. I will finally mention the soundness and completeness theorem for this EL.

For expediency, I will simply state the various validities (and derivable rules) of this logic:

Epistemic Validities:

(**EL1**) $(K\phi \wedge K(\phi \rightarrow \psi)) \rightarrow K\psi$ (Axiom **K**, reflecting Def. 2.1.1(a))

(**EL2**) $K\phi \rightarrow \phi$ (Axiom **T**, reflecting Def. 2.1.1(c))

Epistemic Necessitation Rule:

If ϕ is a local validity, then $K\phi$ (reflecting Def. 2.1.1(b))

Doxastic Validities:

(**EL3**) $(B\phi \wedge B(\phi \rightarrow \psi) \rightarrow B\psi$ (Axiom **K**, reflecting Def. 2.1.2(a))

(**EL4**) $\neg B\perp$ (Axiom **D**, reflecting Def. 2.1.2(c))

Doxastic Necessitation Rule:

If ϕ is a validity, then $B\phi$ (reflecting Def. 2.1.2(b))

Mixed Validities:

(**EL5**) $K\phi \rightarrow B\phi$ (reflecting Prop. 2.2.1)

Take as a final axiom the following:

(**EL6**) The axioms of propositional calculus hold

These observations regarding the syntax of epistemic logic are married to the semantics outlined by the following theorem:

THEOREM 3.5.1 The axioms (**EL1**) to (**EL6**), in addition to the necessitation rules and modus ponens, are *sound and complete* with respect to the semantics put forward in §3.3.2 and §3.4.

Pf. See Kraus and Lehmann (1986). □

I feel I should provide a remark on the soundness and completeness theorem. The soundness and completeness theorem essentially means that all of the of the ideas presented in §2.1.1, §2.1.2, §2.2.1 and §3.3.1 are captured in EL. More over, *soundness* states that the axioms given can be used to reason about all systems satisfying the semantics in §3.3.2 and §3.4. On the other hand, *completeness* states that for anything proposition the axioms fail to deduce, a counter example may be found among the models that conform to §3.3.2 and §3.4.

3.6 Possible Axioms

In this section I move from the core of epistemic logic to statements which may be held more provisionally. I will give a brief philosophical discussion for these statements, which will largely follow the discussion by Hintikka (1962).

(**EL7**) $B\phi \to KB\phi$ (occasionally referred to as *consciousness*)

(**EL8**) $B\phi \to BK\phi$ (occasionally referred to as *positive certainty*)

With regards to the above statements, Hintikka uses some of them in his treatment of Moore's paradox[6]. Hintikka does not find the above propositions entirely necessary in any of the arguments he puts forward, however; he apparently regards them as making strong statements about epistemic logic he is not ready to commit himself to.

Hintikka identifies a common philosophical objection to principles (**EL7**) and (**EL8**). An epistemic alternative to a world s, call it s', is simply a state of affairs that is consistent with all that one knows. One could believe more, less, or different things altogether at this epistemic alternative. Imposing that $B\phi \to KB\phi$ for instance means, in a sense, that in order for a world to be consistent with all that one knows, it must at least be consistent with all that one believes. However, given new evidence one might believe different things, and epistemic alternatives can have evidence that reveal what one believes to be wrong. Analogously, asserting that $B\phi \to BB\phi$ states that for every (doxastic) alternative to ones beliefs one considers, no new evidence could emerge to overturn one's existing beliefs. The statements $B\phi \to BK\phi$ is even stronger still.

In the literature, certain other axioms are popular as well; since Hintikka's writing, they have taken on special titles:

(**EL9**) $K\phi \to KK\phi$ (Axiom **4**, *positive epistemic introspection*)

(**EL10**) $\neg K\phi \to K\neg K\phi$ (Axiom **5**, *negative epistemic introspection*)

Note that analogues of positive and negative introspection exist for doxastic logic as well:

[6]Moore's paradox states "p but I do not believe that p"; Hintikka did not consider the sentence a paradox in his logic, however. Hintikka interpreted the sentence as a falsity under an appropriate interpretation (see Hintikka, 1962; Moore and Shaw, 2005).

(**EL11**) $B\phi \rightarrow BB\phi$ (Axiom **4**, *positive doxastic introspection*)

(**EL12**) $\neg B\phi \rightarrow B\neg B\phi$ (Axiom **5**, *negative doxastic introspection*)

Hintikka heavily endorses positive epistemic introspection, and holds positive doxastic introspection more or less provisionally. Hintikka argues for positive epistemic introspection by endorsing an argument due to Schopenhaur. Hintikka quotes Schopenhaur as saying:

> If your knowing and your knowing that you know are two different things, just try to separate them, and first to know without knowing that you know, then to know that you know, without this knowledge being at the same time knowing (Hintikka, 1962; Schopenhauer, 1889).

Hintikka goes on to demonstrate how he feels Schopenhaur's argument reflects an argument he himself develops, which employs a notion of possible worlds and certain principles he lays out regarding possible worlds. Also, while Hintikka endorses positive epistemic introspection, he rejects arguments *from* introspection. Hintikka holds that one does not necessarily have direct access to their psychology. Hintikka also holds that no sound arguments can be made to show positive *doxastic* introspection. The principles that Hintikka lays governing belief are different somewhat from knowledge, so no analogous argument to Schopenhaur's argument is available for belief.

A note should be made for negative introspection: Hintikka does not think that one could motivate such a proposition. Note that the statement of negative epistemic introspection is equivalent to $P\phi \rightarrow KP\phi$; which is to say something analogous to "If something is possible, as far as the agent knows, then she knows that it is possible as far as she knows." Hintikka writes "It is not true that everybody could come to know the possibility of any fact whatsoever simply by following the consequences of what [s]he already knows" (Hintikka, 1962, p.55).

Despite philosophical objections, however, belief is commonly modeled with negative introspection (see Meyer and van der Hoek, 1995, pp. 68). In fact, a modal logic of knowledge is commonly modeled as **S5** and the modal logic of belief is commonly modeled as **KD45**. However, in the logic I develop, I will demonstrated that negative doxastic introspection is incompatible with other possible propositions which seem more acceptable. I will return to this point in §4.8.

In this essay, I have held all of these principles provisionally. I have done this intentionally; it is entirely clear how they relate to the chain of inclusion I illustrate in §2.2. Certain principles may be more suitable to certain applications; namely, there are clear applications of an **S5** model of knowledge to game theory Fagin et al. (1995); Halpern (1999). I will demonstrate shortly that some of these principles jointly lead to a collapse of belief into knowledge. I hold this as evidence that one should be careful when choosing axioms for systems of knowledge and belief.

3.7 The Collapse of Knowledge and Belief

In this section I investigate the problem of the collapse of knowledge and belief (Halpern, 1996; Lenzen, 1978; van der Hoek, 1991). I already mentioned that computer scientists commonly model knowledge as **S5** and **KD45**. The addition of *consciousness*, axiom (**EL7**), and *positive certainty*, axiom (**EL8**), lead to the collapse of knowledge into belief. I will illustrate this in the following theorem:

THEOREM 3.7.1 Suppose the following axioms hold:

(*I*) All propositional tautologies

(*II*) $(B\phi \wedge B(\phi \rightarrow \psi)) \rightarrow B\psi$

(*III*) $K\phi \rightarrow B\phi$

(*IV*) $\neg K\phi \rightarrow K\neg K\phi$

(*V*) $B\phi \rightarrow BK\phi$

(*VI*) $\neg B\bot$

... then $\vdash B\phi \leftrightarrow K\phi$

Pf. It is instructive to prove the following result:

$$\vdash (B\phi \wedge B\psi) \rightarrow B(\phi \wedge \psi)$$

(*1*) $\vdash (B\phi \wedge B(\phi \rightarrow (\psi \rightarrow (\phi \wedge \psi)))) \rightarrow B(\psi \rightarrow (\phi \wedge \psi))$ (*II*)

(*2*) $\vdash (B\psi \wedge B(\psi \rightarrow (\phi \wedge \psi))) \rightarrow B(\phi \wedge \psi)$ (*II*)

(*3*) $\vdash \phi \rightarrow (\psi \rightarrow (\phi \wedge \psi))$ *tautology*, (*I*)

(*4*) $\vdash B(\phi \rightarrow (\psi \rightarrow (\phi \wedge \psi)))$ (*3*) Nec.

(*5*) $\vdash ((B\phi \wedge B(\phi \rightarrow (\psi \rightarrow (\phi \wedge \psi)))) \rightarrow B(\psi \rightarrow (\phi \wedge \psi))) \rightarrow (((B\psi \wedge B(\psi \rightarrow (\phi \wedge \psi))) \rightarrow B(\phi \wedge \psi)) \rightarrow (B(\phi \rightarrow (\psi \rightarrow (\phi \wedge \psi))) \rightarrow ((B\phi \wedge B\psi) \rightarrow B(\phi \wedge \psi))))$ *tautology*, (*I*)

(*6*) $\vdash ((B\psi \wedge B(\psi \rightarrow (\phi \wedge \psi))) \rightarrow B(\phi \wedge \psi)) \rightarrow (B(\phi \rightarrow (\psi \rightarrow (\phi \wedge \psi))) \rightarrow ((B\phi \wedge B\psi) \rightarrow B(\phi \wedge \psi)))$ (*1*), (*5*) MP

(*7*) $\vdash B(\phi \rightarrow (\psi \rightarrow (\phi \wedge \psi))) \rightarrow ((B\phi \wedge B\psi) \rightarrow B(\phi \wedge \psi))$ (*2*), (*6*) MP

(*8*) $\vdash (B\phi \wedge B\psi) \rightarrow B(\phi \wedge \psi)$ (*4*), (*7*) MP

It is also useful to employ the *hypothetical syllogism*, which states $\vdash \phi \rightarrow \psi, \vdash \psi \rightarrow \chi / \vdash \phi \rightarrow \chi$. The proof is straightforward:

(*1*) $\vdash \phi \rightarrow \psi$ Assumption

(*2*) $\vdash \psi \rightarrow \chi$ Assumption

(*3*) $\vdash (\phi \rightarrow \psi) \rightarrow ((\psi \rightarrow \chi) \rightarrow (\phi \rightarrow \chi))$ *tautology*, (*I*)

(4) $\vdash (\psi \to \chi) \to (\phi \to \chi)$ (1), (3) MP

(5) $\vdash \phi \to \chi$ (2), (4) MP

With those results behind me, I move to proving the result:

(1) $\vdash \neg K\phi \to K\neg K\phi$ (IV)

(2) $\vdash K\neg K\phi \to B\neg K\phi$ (III)

(3) $\vdash \neg K\phi \to B\neg K\phi$ (1), (2) HS

(4) $\vdash (B\neg K\phi \wedge BK\phi) \to B(K\phi \wedge \neg K\phi)$ from above

(5) $\vdash ((B\neg K\phi \wedge BK\phi) \to B(K\phi \wedge \neg K\phi)) \to (\neg B(K\phi \wedge \neg K\phi) \to \neg(B\neg K\phi \wedge BK\phi))$ *tautology*, (I)

(6) $\vdash \neg B(K\phi \wedge \neg K\phi) \to \neg(B\neg K\phi \wedge BK\phi)$ (4), (5) MP

(7) $\vdash \neg B(K\phi \wedge \neg K\phi)$ (VI)

(8) $\vdash \neg(B\neg K\phi \wedge BK\phi)$ (7), (6) MP

(9) $\vdash \neg(B\neg K\phi \wedge BK\phi) \to (B\neg K\phi \to \neg BK\phi)$ *tautology*, (I)

(10) $\vdash B\neg K\phi \to \neg BK\phi$ (8), (9) MP

(11) $\vdash \neg K\phi \to \neg BK\phi$ (2), (10) HS

(12) $\vdash (\neg K\phi \to \neg BK\phi) \to (BK\phi \to K\phi)$ *tautology*, (I)

(13) $\vdash BK\phi \to K\phi$ (11), (12) MP

(14) $\vdash B\phi \to BK\phi$ (V)

(15) $\vdash B\phi \to K\phi$ (13), (14) MP

(16) $\vdash (B\phi \to K\phi) \to ((K\phi \to B\phi) \to (B\phi \leftrightarrow K\phi))$ *tautology*, (I)

(17) $\vdash (K\phi \to B\phi) \to (B\phi \leftrightarrow K\phi)$ (15), (16) MP

(18) $\vdash K\phi \to B\phi$ (III)

(19) $\vdash B\phi \leftrightarrow K\phi$ (18), (17) MP

□

3.8 Closing Remarks

In this section I provide conclusions based on the observations of the last section. I will first remark on the general project of epistemic logic and then remark on the matter of the collapse of belief into knowledge.

I first mention the general project of epistemic logic. Epistemic logic mostly takes a bird's eye view at the dynamics of logic in individuals or groups of individuals. I omitted mention of approaches to epistemic logic that involve common knowledge and common belief, or time indexicals, however such approaches certainly exist (see Fagin et al., 1995; Meyer and van der Hoek, 1995). While there is significant interest in the behavior of knowledge bases, the focus is basically on internal consistency, not correspondence with reality (Halpern, 1997; Halpern and Moses, 1984; van der Hoek and Thijsse, 1994). Discussion of knowledge bases appears to be largely external to epistemic logic itself; epistemic logic seems to be concerned with the dynamics of knowledge and belief, without much regard for how an agents comes to hold beliefs.

The collapse of the belief into knowledge as observed in the previous section has been met with various responses. One response, due to Halpern (1996), is to deny that knowledge entails belief. If knowledge is restricted to propositions consisting only of sentence letters, then the proof in the prvious section does not go through. Halpern feels this solution is adequate, because many of the examples in the philosophical literature explicating how knowledge entails belief (Halpern was namely looking at Lenzen, 1978) do not involve epistemic or doxastic statements, only propositions. Halpern does not claim this is a one and true solution to the problem, only that it is a solution; other authors have offered a variety of other solutions (Kraus and Lehmann, 1986; van der Hoek, 1991; Voorbraak, 1992). Halpern writes that there is no best solution; what system interpretation of knowledge is appropriate depends on application. I agree with this sentiment. However, in codifying the results from §2, I have unearthed yet another way belief may collapse into knowledge, based on the bounds on belief I have articulated there. I will investigate this in the next section.

4 A Logic of *False Lemmas* (FL)

4.1 Introduction

In this section, I investigate the idea of developing a philosophic logic surrounding the sort of reasoning I employed in §2.2. I call this logic *false lemma* logic, because it concerns when one has possibly arrived at a belief through a so called "false lemma," or specious reasoning, and when one has not. The main project of this section is to illustrate a way of codifying the philosophical results I arrived at in Eq. (2.2.5) in a system of logic, and investigate the nature of such semantics. My presentation will follow the presentation I give for epistemic logic EL, however I will provide more analysis and theorems of FL, and provide more details. I will first present the language of my logic and then motivate the semantics philosophically. I will then turn to giving a formal presentation of the semantics, and pay special detail to the accessibility relations I employ. After that, I will present several theorems regarding my semantics, followed by a treatment of possible axioms. I conclude with a presentation of *negative doxastic introspection* and demonstrate how it leads to the collapse of the knowledge and belief relations in FL.

Note that just as in the case of traditional epistemic logic, I follow Meyer and van der Hoek (1995), despite my philosophical loyalties to Hintikka (1962). FL distinguishes itself from other epistemic logics in its definition of the semantics of the R_B relation as well the semantics regarding new relations R_+ and R_-; special detail will be paid to these differences.

4.2 Grammar

In this section, I present the grammar of FL, which is very explicitly an extension of EL, the language of epistemic logic, already presented in §3.2. I will define two grammars in this section, one of them will be used for basies of belief while the other will be used for the language in general. Recall that $\mathbf{P}$ is a set of propositional atoms or constants. I will first give the full blown grammar:

Definition 4.2.1 Let $\mathscr{L}_{\mathsf{FL}}(\mathbf{P})$ be the smallest set such that:

(*a*) $\mathbf{P} \subseteq \mathscr{L}_{\mathsf{FL}}(\mathbf{P})$

(*b*) If $\{\phi, \psi\} \subseteq \mathscr{L}_{\mathsf{FL}}(\mathbf{P})$, then $(\phi \wedge \psi) \in \mathscr{L}_{\mathsf{FL}}(\mathbf{P})$

(*c*) If $\phi \in \mathscr{L}_{\mathsf{FL}}(\mathbf{P})$, then $\{\neg\phi, K\phi, B\phi, \boxplus\phi, \boxminus\phi\} \subseteq \mathscr{L}_{\mathsf{FL}}(\mathbf{P})$

Next, I provide a restriction on this grammar suitable for basies of belief. This restriction will be made apparent in the subsequent discussion.

Definition 4.2.2 Let $\mathscr{L}^*_{\mathsf{FL}}(\mathbf{P})$ be the smallest set such that:

(*a*) $\mathbf{P} \subseteq \mathscr{L}^*_{\mathsf{FL}}(\mathbf{P})$

(*b*) If $\{\phi, \psi\} \subseteq \mathscr{L}^*_{\mathsf{FL}}(\mathbf{P})$, then $(\phi \wedge \psi) \in \mathscr{L}^*_{\mathsf{FL}}(\mathbf{P})$

(*c*) If $\phi \in \mathscr{L}^*_{\mathsf{FL}}(\mathbf{P})$, then $\neg\phi \in \mathscr{L}^*_{\mathsf{FL}}(\mathbf{P})$

The remarks regarding the language of epistemic logic EL apply to this discussion as well; I will use common abbreviations and omit parenthesis. I will intuitively use $\diamondplus\phi$ to denote $\neg\boxplus\neg\phi$, and $\diamondminus\phi$ to denote $\neg\boxminus\neg\phi$. Finally, as before, I will abbreviate $\mathscr{L}_{\mathsf{FL}}(\mathbf{P})$ and $\mathscr{L}^*_{\mathsf{FL}}(\mathbf{P})$ to simply $\mathscr{L}_{\mathsf{FL}}$ and $\mathscr{L}^*_{\mathsf{FL}}$.

4.3 Semantics

In this section I provide the philosophical motivation and semantics for FL. The semantics for the R_B relation will be defined in terms of belief basis which I will carry around in my logic. The relations R_+ and R_- will allow for access of *grounded* and *specious* subsets of the agent's belief basis. I will first give an informal philosophical explanation and then a formal defintion.

4.3.1 Philosophical Motivation for Semantics

In this section I attempt to provide a philosophical motivation for the semantics of FL. I will first provide some explanation regarding the basic idea behind my interpretation of the doxastic accessibility relation, which is given by R_B. I will then talk about accessibility relations which will allow for to access grounded and specious subsets of the agent's belief base, namely R_+, and R_-. Finally, I talk about my conception of possible worlds.

I intend to emulate the semantics of traditional epistemic logic, largely, as lain out in §3.3.1. The most significant addition being that I will be concerned with belief basis as well, reflecting the definition in §2.2.2. *The idea of this logic is that a belief base $\mathscr{B}$ is generated by a belief basis* $\mathfrak{B}$. I intend for a belief basis to *determine* a set of doxastic alternatives for an agent at a particular possible world; a reader should note that this is radically different from ordinary epistemic logic. In FL the doxastic accessibility relation is, in a sense, *functional* of an agent's belief basies. I feel a semantics that will be tracking belief will then have to track an agent's belief basies as well; to this end I will consider an agent's belief basis $\mathfrak{B}$ to be a fragment of a possible world, however it will be a fragment of a possible world which will not bear on truth of sentence letters in $\mathbf{P}$. In this respect, I am essentially making a dualistic assumption: an agent's beliefs are separate from the rest of the model she occupies. Thus, the possible worlds I have in mind are more complicated than usual Kripke semantics.

With this added level of complexity, I intend to mirror my philosophical analysis I have previously developed. Namely, I build reasoning around the observation of §2.2; which was that, depending on the properties ϕ has or not with respect to certain subsets of an agent's belief basis, it can be either be considered knowledge or ruled out. Recall that I argued that if ϕ is known then ϕ is *possibly grounded*, which I asserted in Prop. 2.2.3. Also recall that I argued that if a proposition is *fully grounded* then it is knowledge, as I wrote in §2.2.4. To find possibly grounded beliefs, I am interested in carving up an agent's belief basis; to this end I introduce R_+, which will allow for access to grounded subsets of an agent's belief basis. I will also be interested in investigating R_-, which will allow for to access *specious* subsets of an agent's beliefs. I will not introduce a *fully grounded* operator, for the notion of what it means for a statement to be fully grounded may be characterized in terms of R_+ and R_-.

In order satisfy my design criteria, my logic will have to differ from ordinary epistemic logic somewhat. First and foremost, I represent worlds as ordered pairs of the form $\langle b, \mathfrak{B}\rangle$, where b denotes the the state of affairs and $\mathfrak{B}$ denotes the agent's belief basis. Specifically, in intend for b to determine the state of affairs by determining the truth of sentence letters. I also intend for $\mathfrak{B}$ to determine the doxastic, grounded, and specious relations behave (note that I intentionally leave the epistemic relation open, as it is the object of study in this essay). In a sense, the semantics I envision has two dimensions; one for states of affairs, and one for beliefs basies. This second dimension determines various relations I am interested in investigating by allowing my semantics to *talk about itself*, in a sense. I will demonstrate how my semantics works shortly.

4.3.2 Basic Semantics

I submit the following definition of a *modified Kripke structure*:

Definition 4.3.1 A *modified Kripke structure* $\mathbb{M}$ is a tuple $\langle \Omega, \pi, R_K, R_B, R_+, R_- \rangle$ where:

(*i*) $\Omega \subseteq S \times \mathscr{P}(\mathscr{L}^*_{\mathsf{FL}}(\mathbf{P}))$ where S is a non-empty set of *states*

(*ii*) $\pi : \Omega \to (\mathbf{P} \to \{\mathsf{True}, \mathsf{False}\})$ is a truth assignment to a proposition at a particular world, such that

$$\text{If } \langle b, \mathfrak{B} \rangle, \langle b, \mathfrak{C} \rangle \in \Omega, \text{ then } \pi(\langle b, \mathfrak{B} \rangle)(p) = \pi(\langle b, \mathfrak{C} \rangle)(p) \text{ for all } p \in \mathbf{P}$$

(*iii*) $R_K \subseteq \Omega \times \Omega$, $R_B \subseteq \Omega \times \Omega$, $R_+ \subseteq \Omega \times \Omega$, and $R_- \subseteq \Omega \times \Omega$

The relation $(\mathbb{M}, \langle b, \mathfrak{B} \rangle) \models \phi$ (that is, ϕ is satisfied by world $\langle b, \mathfrak{B} \rangle$ in model $\mathbb{M}$) is defined inductively as follows:

$$\begin{aligned}
(\mathbb{M}, \langle b, \mathfrak{B} \rangle) \models p &\iff \pi(s)(p) = \mathsf{True} \text{ where } p \in \mathbf{P} \\
(\mathbb{M}, \langle b, \mathfrak{B} \rangle) \models \phi \wedge \psi &\iff (\mathbb{M}, \langle s, \mathfrak{B} \rangle) \models \phi \text{ and } (\mathbb{M}, \langle s, \mathfrak{B} \rangle) \models \psi \\
(\mathbb{M}, \langle b, \mathfrak{B} \rangle) \models \neg\phi &\iff (\mathbb{M}, \langle b, \mathfrak{B} \rangle) \not\models \phi \\
(\mathbb{M}, \langle b, \mathfrak{B} \rangle) \models \Box\phi &\iff (\mathbb{M}, \langle x, \mathfrak{X} \rangle) \models \phi \text{ for all } \langle x, \mathfrak{X} \rangle \text{ with } R_\Box(\langle b, \mathfrak{B} \rangle, \langle x, \mathfrak{X} \rangle)
\end{aligned}$$

... where $\Box$ is $K, B, \boxplus$, or $\boxminus$

The notation used in Def. 3.3.2 for EL, which regards the notions of *validities, local validities* and what it means for a proposition or set to be *satisfied*, will be reused for FL with essentially no loss of meaning.

Note that in the semantics outlined that, so far, the only role a belief basis $\mathfrak{B}$ has is a descriptor of a possible world. One could safely ignore them and carry on with ordinary epistemic logic as usual. Despite this, I hold it plays a role in the nature of the accessibility relations.

4.4 Characterizing Accessibility Relations

In this section I will characterize the accessibility relations in the semantics of FL. Because my logic is somewhat more exotic than ordinary epistemic logic, special attention shall be paid to these presentations.

4.4.1 The R_B Accessibility Relation

In this section I define the semantics for the R_B relation. The semantics for R_B will be shown to depend completely on the belief basies I have incorporated into the possible world semantics above. I will conclude this discussion with a paradox that emerges if one permits for $\mathscr{L}^*_{\mathsf{FL}}$ to be less restrictive.

I will start by presenting the definition of the R_B relation. The essential idea I have is that I want $R_B(\langle b, \mathfrak{B} \rangle, \langle c, \mathfrak{C} \rangle)$ to be equivalent to saying that $\mathbb{M}, \langle c, \mathfrak{C} \rangle \models \mathfrak{B}$. In other words, another world $\langle c, \mathfrak{C}\rangle$ is considered doxastically accessible from $\langle b, \mathfrak{B}\rangle$ if and only if $\langle c, \mathfrak{C}\rangle$ makes true all the statements in $\mathfrak{B}$. In FL, an doxastically accessible world $\langle c, \mathfrak{C}\rangle$ is literally a realization of the beliefs in another world. It can be seen then that $\mathfrak{B}$ therefore *generates* an agents beliefs. These ideas are codified in the following proposition:

PROPOSITION 4.4.1 $R_B(\langle b, \mathfrak{B}\rangle, \langle c, \mathfrak{C}\rangle)$ if and only if $(\mathbb{M}, \langle c, \mathfrak{C}\rangle) \models \mathfrak{B}$

I now explain a constraint on $\mathscr{L}^*_{\mathsf{FL}}$ I have made. One might think that $\mathscr{L}^*_{\mathsf{FL}}$ could be more complex than it is, and perhaps include formulae regarding belief. I am reluctant to do so, as permitting this would lead to potential paradoxes. The following example illustrates this:

> EXAMPLE 4.4.1 Let $\mathbf{P} = \{q\}$. Construct the following modified Kripke structure
>
> $$\mathbb{M} = \langle \Omega, \pi, R_K, R_B, R_+, R_- \rangle$$
>
> Let $\Omega = \{\langle \omega, \{B(q)\}\rangle\}$, $\pi(\langle \omega, \{B(q)\}\rangle)(q) = \mathsf{False}$, and relations take any value, provided the restriction of Prop. 4.4.1.
>
> Does $(\mathbb{M}, \langle \omega, \{B(q)\}\rangle) \models B(q)$?
>
> Suppose $(\mathbb{M}, \langle \omega, \{B(q)\}\rangle) \not\models B(q)$, then for any $\phi \in \mathscr{L}_{\mathsf{FL}}$, it would be true for every world $\langle x, \mathfrak{X}\rangle$ such that $R_B(\langle \omega, \{B(q)\}\rangle, \langle x, \mathfrak{X}\rangle)$ vacuously, which would mean that $(\mathbb{M}, \langle \omega, \{B(q)\}\rangle) \models B(q)$, which would be a contradiction.
>
> Suppose $(\mathbb{M}, \langle \omega, \{B(q)\}\rangle) \models B(q)$, then $R_B(\langle \omega, \{B(q)\}\rangle, \langle \omega, \{B(q)\}\rangle)$. This would mean that q is true for all worlds $\langle x, \mathfrak{X}\rangle$ such that $R_B(\langle \omega, \{B(q)\}\rangle, \langle x, \mathfrak{X}\rangle)$, in particular, $(\mathbb{M}, \langle \omega, \{B(q)\}\rangle) \models q$.
> However, $\pi(\langle \omega, \{B(q)\}\rangle)(q) = \mathsf{False}$, which means that
> $(\mathbb{M}, \langle \omega, \{B(q)\}\rangle) \not\models q$, which is a contradiction.

... thus, in this epistemic logic, for an agent to have direct experience or direct evidence regarding its own beliefs is banned, to avoid such degenerate scenarios as the one above.

4.4.2 ***Horizontal*** & ***Vertical*** Accessibility Relations

In this section I provide vocabulary to talk about novel properties modal operators can have. Accompanying this vocabulary, I will present a theorem regarding these concepts. I will conclude with a discussion of how this novel vocabulary allows me to characterize other accessibility relations I am interested in.

The new vocabulary is summarized in the following definition:

DEFINITION 4.4.1 Define the following terms:

(a) If for every $\langle b, \mathfrak{B}\rangle$ and $\langle c, \mathfrak{C}\rangle$ in $\mathbb{M}$, if $R_\Box(\langle b, \mathfrak{B}\rangle, \langle c, \mathfrak{C}\rangle)$ then $\mathfrak{B} \subseteq \mathfrak{C}$, then $R_\Box$ is called *nondecreasing* in $\mathbb{M}$.

(b) If for every $\langle b, \mathfrak{B}\rangle$ and $\langle c, \mathfrak{C}\rangle$ in $\mathbb{M}$, if $R_\Box(\langle b, \mathfrak{B}\rangle, \langle c, \mathfrak{C}\rangle)$ then $\mathfrak{B} \supseteq \mathfrak{C}$, then $R_\Box$ is called *nonincreasing* in $\mathbb{M}$.

(c) If for every $\langle b, \mathfrak{B}\rangle$ and $\langle c, \mathfrak{C}\rangle$ in $\mathbb{M}$, if $R_\Box(\langle b, \mathfrak{B}\rangle, \langle c, \mathfrak{C}\rangle)$ then $\mathfrak{B} = \mathfrak{C}$, then $R_\Box$ is called *horizontal* in $\mathbb{M}$.

(d) If for every $\langle b, \mathfrak{B}\rangle$ and $\langle b, \mathfrak{C}\rangle$ in $\mathbb{M}$, if $R_\Box(\langle b, \mathfrak{B}\rangle, \langle c, \mathfrak{C}\rangle)$ then $b = c$, then $R_\Box$ is called *vertical* in $\mathbb{M}$.

...where $\Box$ is $K, B, \boxplus, \boxminus$.

With this vocabulary defined, I can motivate my idea behind the various modal operators I have introduced. A *nondecreasing* accessibility relation accesses other possible states where nothing has been removed from the knowledge basis. Likewise, a *nonincreasing* accessibility relation accesses other possible states where nothing has been added. A *horizontal* accessibility relation accesses other possible states but keeps the belief basis constant. A *vertical* accessibility relation accesses other belief basis but keeps the state constant.

Some notes regarding these concepts should be made. I mainly focus on vertical, nondecreasing relations in my analysis, when characterizing R_+ and R_-. I mention nonincreasing relations more rarely; specifically, I will talk about what happens when R_B is nonincreasing when I discuss *negative doxastic introspection*. Finally, while I have introduced the term *horizontal* in the above definition, I do not find use for it in my analysis. I present it here largely because its definition follows naturally from the notions of *nonincreasing* and *nondecreasing*.

With this definition in mind, I derive the following theorem:

THEOREM 4.4.1 If $R_\Box$ is *nonincreasing* in $\mathbb{M}$, then $\mathbb{M} \models B\phi \to \Box B\phi$. If $R_\Box$ is *nondecreasing* in $\mathbb{M}$, then $\mathbb{M} \models \neg B\phi \to \Box\neg B\phi$, where $\Box$ is $K, B, \boxplus$ and $\boxminus$.

Pf. The proof follows from two lemmas, Lemmas 4.4.1 and 4.4.2. $\Box$

LEMMA 4.4.1 If $\mathfrak{B} \subseteq \mathfrak{X}$ for all $\mathfrak{X}$ such that $R_\Box(\langle b, \mathfrak{B}\rangle, \langle x, \mathfrak{X}\rangle)$ for some x, then for all $\phi \in \mathscr{L}_{\mathsf{FL}}$, $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models B\phi \to \Box B\phi$ where $\Box$ is either $K, B, \boxplus, \boxminus$.

Pf. Assume that $\mathfrak{B} \subseteq \mathfrak{X}$ for all $\mathfrak{X}$ such that $R_\Box(\langle b, \mathfrak{B}\rangle, \langle x, \mathfrak{X}\rangle)$ for some x.

If $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \not\models B\phi$, then the result is vacuously true. So assume that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models B\phi$; it suffices to show that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \Box B\phi$.

To this end, let the pair $\langle c, \mathfrak{C}\rangle$ be such that $R_\Box(\langle b, \mathfrak{B}\rangle, \langle c, \mathfrak{C}\rangle)$. To show $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \Box B\phi$, it suffices to show $(\mathbb{M}, \langle c, \mathfrak{C}\rangle) \models B\phi$. Let $\langle d, \mathfrak{D}\rangle$ be such that $R_B(\langle c, \mathfrak{C}\rangle, \langle d, \mathfrak{D}\rangle)$. To show $(\mathbb{M}, \langle c, \mathfrak{C}\rangle) \models B\phi$, it suffices to show $(\mathbb{M}, \langle d, \mathfrak{D}\rangle) \models \phi$.

I argue that $R_B(\langle b, \mathfrak{B}\rangle, \langle d, \mathfrak{D}\rangle)$. To show this, it suffices to show, by Prop. 4.4.1, that $(\mathbb{M}, \langle d, \mathfrak{D}\rangle) \models \mathfrak{B}$. Recall that $\mathfrak{B} \subseteq \mathfrak{C}$ by assumption and $R_B(\langle c, \mathfrak{C}\rangle, \langle d, \mathfrak{D}\rangle)$. Therefore, $(\mathbb{M}, \langle d, \mathfrak{D}\rangle) \models \mathfrak{C}$ by Prop. 4.4.1. That is, by Def. 3.3.2(vi), $(\mathbb{M}, \langle d, \mathfrak{D}\rangle) \models \psi$ for all $\psi \in \mathfrak{C}$. However, by assumption I have that $\mathfrak{B} \subseteq \mathfrak{C}$, therefore $(\mathbb{M}, \langle d, \mathfrak{D}\rangle) \models \psi$ for all $\psi \in \mathfrak{B}$. This just means, by definition, that $(\mathbb{M}, \langle d, \mathfrak{D}\rangle) \models \mathfrak{B}$. Therefore, $R_B(\langle b, \mathfrak{B}\rangle, \langle d, \mathfrak{D}\rangle)$.

From my assumption that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models B\phi$ and $R_B(\langle b, \mathfrak{B}\rangle, \langle d, \mathfrak{D}\rangle)$, then $(\mathbb{M}, \langle d, \mathfrak{D}\rangle) \models \phi$ by Prop. 4.4.1. Thus the proof is complete. $\square$

Lemma 4.4.2 If $\mathfrak{B} \supseteq \mathfrak{X}$ for all $\mathfrak{X}$ such that $R_\square(\langle b, \mathfrak{B}\rangle, \langle x, \mathfrak{X}\rangle)$ for some x, then for all $\phi \in \mathscr{L}_{\mathsf{FL}}$, $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \neg B\phi \to \square\neg B\phi$, where $\square$ is either $K, B, \boxplus, \boxminus$.

Pf. The proof is very similar to the previous lemma. Assume that $\mathfrak{B} \supseteq \mathfrak{X}$ for all $\mathfrak{X}$ such that $R_\square(\langle b, \mathfrak{B}\rangle, \langle x, \mathfrak{X}\rangle)$ for some x. Also assume $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \neg B\phi$; it suffices to show that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \square\neg B\phi$.

To this end, let the pair $\langle c, \mathfrak{C}\rangle$ be such that $R_\square(\langle b, \mathfrak{B}\rangle, \langle c, \mathfrak{C}\rangle)$. To show $(\mathbb{M}, \langle c, \mathfrak{C}\rangle) \models \neg B\phi$, it suffices to find some $\langle d, \mathfrak{D}\rangle$ such that $R_B(\langle c, \mathfrak{C}\rangle, \langle d, \mathfrak{D}\rangle)$ where $(\mathbb{M}, \langle d, \mathfrak{D}\rangle) \models \neg\phi$.

I have $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \neg B\phi$ by assumption; this means that there is a world $\langle f, \mathfrak{F}\rangle$ such that $R_B(\langle b, \mathfrak{B}\rangle, \langle f, \mathfrak{F}\rangle)$ and $(\mathbb{M}, \langle f, \mathfrak{F}\rangle) \models \neg\phi$. I argue that $R_B(\langle c, \mathfrak{C}\rangle, \langle f, \mathfrak{F}\rangle)$.

By Prop. 4.4.1, $R_B(\langle b, \mathfrak{B}\rangle, \langle f, \mathfrak{F}\rangle)$ if and only if $(\mathbb{M}, \langle f, \mathfrak{F}\rangle) \models \mathfrak{B}$. However, because $\mathfrak{C} \subseteq \mathfrak{B}$, then $(\mathbb{M}, \langle f, \mathfrak{F}\rangle) \models \mathfrak{C}$. Therefore, $R_B(\langle c, \mathfrak{C}\rangle, \langle f, \mathfrak{F}\rangle)$, and furthermore, $(\mathbb{M}, \langle f, \mathfrak{F}\rangle) \models \neg\phi$. As asserted above, this suffices to show that $(\mathbb{M}, \langle c, \mathfrak{C}\rangle) \models \neg B\phi$, which completes the proof. $\square$

It should be noted that the converses of the two parts of Thrm. 4.4.1 do not hold. In fact, the same counter example serves to refute the converses of both parts of the theorem:

> Example 4.4.2 Let $\mathbf{P} = \{p, q\}$. Construct the following modified Kripke structure
>
> $$\mathbb{M} = \langle \Omega, \pi, R_K, R_B, R_+, R_-\rangle$$
>
> Let $\Omega = \{\omega\} \times \{\mathfrak{A}, \mathfrak{B}\}$ where $\mathfrak{A} = \{p, q\}$ and $\mathfrak{B} = \{p \wedge q\}$, $\pi(\langle \omega, \mathfrak{A}\rangle)(x) = \pi(\langle \omega, \mathfrak{B}\rangle)(x) =$ True for all $x \in \mathbf{P}$, and let
>
> $$R_B = \{\langle\langle \omega, \mathfrak{A}\rangle, \langle \omega, \mathfrak{A}\rangle\rangle, \langle\langle \omega, \mathfrak{A}\rangle, \langle \omega, \mathfrak{B}\rangle\rangle, \\ \langle\langle \omega, \mathfrak{B}\rangle, \langle \omega, \mathfrak{A}\rangle\rangle, \langle\langle \omega, \mathfrak{B}\rangle, \langle \omega, \mathfrak{B}\rangle\rangle\}$$
>
> ... it can be seen that $(\mathbb{M}, \langle \omega, \mathfrak{A}\rangle) \models Bp \to BBp$ and $\mathfrak{A} \cap \mathfrak{B} = \varnothing$.

Thrm. 4.4.1 gives me the proper means to talk about nature of accessibility relations in this logic. For instance to return to R_B for a moment, it should be noted that if R_B is *nondecreasing* then it is *positively introspective*, and if it is *nonincreasing* then it is *negatively introspective*, reflecting the characterization of epistemic and doxastic notions put forward in §3.6.

4.4.3 The R_+ Accessibility Relation

In this section I present the R_+ operator. It is in the introduction of the notion of the R_+ operator and the R_- operator that FL distinguishes itself, they are the key to expressing the bounds on belief I endorsed in Prop. ??. I will present a philosophical interpretation of the $\oplus$ operator which is governed by R_+, and conclude this discussion with a demonstration of how extension of the semantics leads to paradoxes.

The following proposition characterizes R_+:

PROPOSITION 4.4.2 $R_+(\langle b, \mathfrak{B}\rangle, \langle c, \mathfrak{C}\rangle)$ if and only if

(*a*) $\langle c, \mathfrak{C}\rangle \in \Omega$

(*b*) $b = c$

(*c*) $\mathfrak{C} \subseteq \mathfrak{B}$

(*d*) $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mathfrak{C}$

This operator requires a word or two in English to properly characterize. First of all, it should be noted that R_+ is a *verticle* and *nonincreasing*, as per the Defs . 4.4.1(d) and 4.4.1(b). Intuitively, R_+ accesses only *grounded* subsets of an agent's belief basis, as per Def. 2.2.4(a), while keeping the state the system is in constant (that is, R_+ can only access worlds where the sentence letters have the same truth value).

A remark about the philosophical interpretation of this operator would be useful. When I write $\oplus B\phi$, then that means that the agent has a belief in ϕ for which she can make an argument which is possible grounded; that is to say that ϕ follows from a subset of things she believes, all of which are True. I literally intend the sentence $\oplus B\phi$ to be read as "The agent has a possibly grounded belief that ϕ."

Finally, I should mention paradoxes I have run into in this investigation. As was the case with B, if one lets $\boxplus$ into the language $\mathscr{L}^*_{\mathsf{FL}}$ then paradoxes emerge. To illustrate this, I provide the following example:

EXAMPLE 4.4.3 Let $\mathbf{P} = \{q\}$. Construct the following modified Kripke structure

$$\mathbb{M} = \langle \Omega, \pi, R_K, R_B, R_+, R_- \rangle$$

Let $\Omega = \{\langle \omega, \{\boxplus\neg q\}\rangle\}$ where $\pi(\langle \omega, \{\boxplus\neg q\}\rangle)(q) = \mathsf{True}$, and the other relations satisfy the constraints so far imposed on them.

Does $(\mathbb{M}, \langle \omega, \{\boxplus\neg q\}\rangle) \models \boxplus\neg q$?

Suppose not, then that would mean that $(\mathbb{M}, \langle \omega, \{\boxplus\neg q\}\rangle) \models \oplus q$, that is to say there is some world $\langle b, \mathfrak{C}\rangle$ such that $\mathfrak{C} \subseteq \mathfrak{B}$ and $(\mathbb{M}, \langle \omega, \{\boxplus\neg q\}\rangle) \models \mathfrak{C}$ where $(\mathbb{M}, \langle \omega, \mathfrak{C}\rangle) \models \neg q$. There's only the one subset, and the one world, so the choice is limited to $\mathfrak{C} = \{\boxplus\neg q\}$. Therefore, I arrive at $(\mathbb{M}, \langle \omega, \{\boxplus\neg q\}\rangle) \models \boxplus\neg q \wedge \neg \boxplus \neg q$, which cannot happen.

Suppose that $(\mathbb{M}, \langle \omega, \{\boxplus\neg q\}\rangle) \models \boxplus\neg q$ is the case then. That would mean

$$R_+(\langle \omega, \{\boxplus\neg q\}\rangle, \langle \omega, \{\boxplus\neg q\}\rangle)$$

...which would mean that $(\mathbb{M}, \langle \omega, \{\boxplus\neg q\}\rangle) \models \neg q$, which cannot happen.

4.4.4 The R_- Accessibility Relation

In this section I present the R_- operator, with some motivation. I will then give a philosophical reading of the induced $\diamondminus$ operator, and finally show how letting statements regarding $\boxminus$ into $\mathscr{L}^*_{\mathsf{FL}}$ also leads to a final paradox which I will illustrate.

The definition of R_- is basically the same as the definition of R_+ with one alteration:

Proposition 4.4.3 $R_-(\langle b, \mathfrak{B}\rangle, \langle c, \mathfrak{C}\rangle)$ if and only if

(*a*) $\langle c, \mathfrak{C}\rangle \in \Omega$

(*b*) $b = c$

(*c*) $\mathfrak{C} \subseteq \mathfrak{B}$

(*d*) $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \not\models \mathfrak{C}$

I will provide a brief word on the R_- accessibility relation. It can be seen that R_- is very similar to the R_+ accessibility relation. The only difference between the two is that R_- accesses *specious* subsets of an agent's belief basis only, in contrast to R_+ which only accesses *grounded* subsets. Neither of them affect the state of affairs is in, in terms of the truth and falsity of sentence letters. Also, just as R_+ is *vertical* and *nondecreasing*, so too in R_-.

Just as was the case for $\diamondplus$, I intend for $\diamondminus$ to have a philosophical reading. I hold that $\diamondminus B\phi$ reads "The agent has a possibly specious belief that ϕ." In addition, it makes sense to read $\diamondminus(B\phi \wedge \diamondplus B\phi)$ as saying "While the agent has a possibly specious belief that ϕ, they can fix up their reasoning to have possibly grounded belief that ϕ." As will be demonstrated, much if not all philosophical intuition regarding possibly grounded and possibly specious belief may be read using the logic I extend.

Finally, I have one last paradox to present regarding my restrictions I have put $\mathscr{L}^*_{\mathsf{FL}}$.

Example 4.4.4 Let $\mathbf{P} = \{q\}$. Construct the following modified Kripke structure

$$\mathbb{M} = \langle \Omega, \pi, R_K, R_B, R_+, R_- \rangle$$

Let $\Omega = \{\langle \omega, \{\diamondminus q\}\rangle\}$ where $\pi(\langle \omega, \{\diamondminus q\}\rangle)(q) = \mathsf{False}$, and the other relations satisfy the constraints so far imposed on them.

Does $(\mathbb{M}, \langle \omega, \{\diamondminus q\}\rangle) \models \diamondminus q$?

Suppose not, then that would mean that $(\mathbb{M}, \langle \omega, \{\diamondminus q\}\rangle) \not\models \diamondminus q$, which means that $R_-(\langle \omega, \{\diamondminus q\}\rangle, \langle \omega, \{\diamondminus q\}\rangle)$, and therefore $(\mathbb{M}, \langle \omega, \{\diamondminus q\}\rangle) \models q$, which cannot happen.

Suppose it is the case that $(\mathbb{M}, \langle \omega, \{\diamondminus q\}\rangle) \models \diamondminus q$, that is to say there is some world $\langle b, \mathfrak{C}\rangle$ such that $\mathfrak{C} \subseteq \mathfrak{B}$ and $(\mathbb{M}, \langle \omega, \{\diamondminus q\}\rangle) \not\models \mathfrak{C}$. Just as before, there's only the one subset, and the one world, so the choice is limited to $\mathfrak{C} = \{\diamondminus q\}$. Therefore, I arrive at $(\mathbb{M}, \langle \omega, \{\boxplus \neg q\}\rangle) \models \diamondminus q \wedge \neg \diamondminus q$, which cannot happen.

4.4.5 The R_K Accessibility Relation

I have intentionally left the R_K relation wide open; this is because philosophically, when I go to retrieve the concept of what it means to know something in terms of belief, I do not come up with anything definitive. I would consider allowing K modal operators in $\mathscr{L}^*_{\mathsf{FL}}$, and they would undoubtedly result in paradoxes as was the case with the other modal operators, however analysis proceeds much more smoothly if they are not permitted. It might be desirable to ensure that R_K is *reflexive* as in the case of epistemic logic; I am happier to let the truth of R_K be a *consequence* of propositions I will suggest, intended to reflect philosophical observations I have made.

A remark about the restrictions I have made on $\mathscr{L}^*_{\mathsf{FL}}$ regarding the K modal operator should be made. While no paradoxes immediately emerge, it seems philosophically strange to allow for an agent to have direct experience of one's own *knowledge*. Therefore, I am happy to let psychology, experience and circumstance govern knowledge, at least in this model — it is not something one ever experiences.

As a last remark, if knowledge is *nondecreasing*, then axiom (**EL7**) holds. Recall that (**EL7**) states that $B\phi \to KB\phi$. This mirrors Hintikka's remarks rather nicely — if it is the case that at epistemic alternative worlds that one has the same basis for their beliefs as this world, then if one believes something in this world they will believe it in the epistemic alternative world.

4.5 Consequences and Validities of FL

In this section I present consequences and validities of FL. I will first give some preliminary results, which I present without proof. Throughout this section, I will endeavor to give philosophical commentary of my results after I derive them.

4.6 Preliminary Results

In this section I state without proof several truths in FL, summarized in the following proposition. This proposition is mostly proved verbatim in Meyer and van der Hoek (1995), and likely any other textbook on modal logic:

PROPOSITION 4.6.1

(*i*) If ϕ is an instance of a propositional tautology, $\models \phi$

(*ii*) $\models (\Box\phi \wedge \Box(\phi \to \psi)) \to \Box\psi$

(*iii*) If $\mathbb{M} \models \phi$ then $\mathbb{M} \models \Box\phi$

(*iv*) If $\mathbb{M} \models \phi$, and $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \neg\Box\bot$ then $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \Diamond\phi$

(*v*) If $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \phi$ and $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \phi \to \psi$, then $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \psi$

(*vi*) $\models (\Box\neg\phi \to \neg\phi) \leftrightarrow (\phi \to \Diamond\phi)$

…where $\Box$ is $K, B, \boxplus$ or $\boxminus$, and $\Diamond$ is $P, C, \diamondplus$ or $\diamondminus$.

I have some philosophical remarks regarding the above proposition. I hold that Prop. 4.6.1(iii) states that the agent knows and believes *background knowledge*, reflecting Defs. 2.1.1(b) and 2.1.2(b) respectively, when $\Box$ is K and B respectively. Also, I hold that Prop. 4.6.1(ii) states that K and B are *closed under material implication*, just like in Defs. 2.1.1(a) and 2.1.2(a) respectively. Finally, an astute philosopher will note that Prop. 4.6.1(v) just states that *modus ponens* is valid.

I will make use of these results freely. Many of these results are technical results which I found necessary to cite or cite implicitly the proofs I provide.

4.6.1 Validities

In this section, I state several validities. After stating these results, I will stop to provide a philosophical interpretation of what they say.

Theorem 4.6.1 *The Belief Theorem*

$$\models B\phi \leftrightarrow (\diamondplus B\phi \vee \diamondminus B\phi)$$

Pf. Let $\langle b, \mathfrak{B}\rangle$ be any world in any model $\mathbb{M}$. It suffices to show that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models B\phi \leftrightarrow (\diamondplus B\phi \vee \diamondminus B\phi)$. To prove that statement, it suffices to show two directions.

$(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models (\diamondplus B\phi \vee \diamondminus B\phi) \to B\phi$:

Assume $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondplus B\phi \vee \diamondminus B\phi$. It suffices to show both $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondplus B\phi \to B\phi$ and $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondminus B\phi \to B\phi$.

Without loss of generality, consider the case of $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondplus B\phi \to B\phi$. Recall that $\diamondplus\psi$ is just shorthand for $\neg \boxplus \neg\psi$ as per §4.2. Therefore, it suffices to show that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \neg \boxplus \neg B\phi \to B\phi$.

Note that $\models (\neg B\phi \to \boxplus\neg B\phi) \to (\neg \boxplus \neg B\phi \to B\phi)$ is true because the proposition is a tautology. Therefore, to show $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \neg\boxplus\neg B\phi \to B\phi$ it suffices to show $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \neg B\phi \to \boxplus\neg B\phi$. However, R_+ is defined to be *nonincreasing*, thus by Thrm. 4.4.1, $\models \neg B\phi \to \boxplus\neg B\phi$. Thus, this subconclusion is true.

$(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models B\phi \to (\diamondplus B\phi \vee \diamondminus B\phi)$:

Assume $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models B\phi$. Either $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mathfrak{B}$ or $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \not\models \mathfrak{B}$. It suffices to show that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondplus B\phi \vee \diamondminus B\phi$ in either case.

If $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mathfrak{B}$, then $R_+(\langle b, \mathfrak{B}\rangle, \langle b, \mathfrak{B}\rangle)$. As I assumed $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mathfrak{B}$, and therefore $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondplus B\phi$. Thus $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondplus B\phi \vee \diamondminus B\phi$.

If $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \not\models \mathfrak{B}$, then $R_-(\langle b, \mathfrak{B}\rangle, \langle b, \mathfrak{B}\rangle)$, Therefore $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondplus B\phi$, and thus

$$(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondplus B\phi \vee \diamondminus B\phi$$

…then, in either case, $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondplus B\phi \vee \diamondminus B\phi$.

$\square$

To provide an interpretation of the above theorem, $\models B\phi \leftrightarrow (\diamondplus B\phi \vee \diamondminus B\phi)$ states two things: (1) "whenever the agent has a belief in ϕ, it is either possibly grounded or possibly specious" and (2) "whenever the agent has a possibly grounded or possibly specious belief in ϕ, then they believe ϕ."

Theorem 4.6.2 *Justification Theorem*

$$\models \diamondplus B\phi \rightarrow \phi$$

Pf. It suffices to show that if $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondplus B\phi$ then $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \phi$. If $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondplus B\phi$ then there is some $\mathfrak{C} \subseteq \mathfrak{B}$ where $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mathfrak{C}$ and $(\mathbb{M}, \langle b, \mathfrak{C}\rangle) \models B\phi$. However, by definition, if $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mathfrak{C}$ then $R_B(\langle b, \mathfrak{C}\rangle, \langle b, \mathfrak{B}\rangle)$; therefore if $(\mathbb{M}, \langle b, \mathfrak{C}\rangle) \models B\phi$ then $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \phi$, which suffices to prove the theorem. $\square$

This theorem says that if the agent has a possibly grounded belief that ϕ, then ϕ is true. This corresponds to the intuition that an agent has a possibly grounded belief in a statement ϕ if and only if they can give a sound argument for ϕ. As many philosophers will agree, the conclusions of sound arguments are True, which is all that this theorem says.

Lemma 4.6.1 $\pm$ *Lemma*

Suppose $\mathfrak{X} \subseteq \mathscr{L}^*_{\mathsf{FL}}$. Then for all $\langle b, \mathfrak{B}\rangle, \langle b, \mathfrak{C}\rangle \in \Omega$,

$$(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mathfrak{X} \iff (\mathbb{M}, \langle b, \mathfrak{C}\rangle) \models \mathfrak{X}$$

Pf. The proof proceeds, without loss of generality, in only one direction.

I want to show if $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mathfrak{X}$ then $(\mathbb{M}, \langle b, \mathfrak{C}\rangle) \models \mathfrak{X}$. The proof proceeds by induction on complexity of $\phi \in \mathfrak{X}$.

First, consider the case of sentence letters $p \in \mathbf{P}$. I am given that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models p$, and therefore $\pi(\langle b, \mathfrak{B}\rangle)(p) = \mathsf{True}$. However, by definition $\pi(\langle b, \mathfrak{B}\rangle)(p) = \pi(\langle b, \mathfrak{C}\rangle)(p)$, therefore $(\mathbb{M}, \langle b, \mathfrak{C}\rangle) \models p$.

Second, consider the case of negations of sentence letters, $\neg p \in \mathbf{P}$. I am given that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \neg p$, and therefore $\pi(\langle b, \mathfrak{B}\rangle)(p) = \mathsf{False}$. However, by definition $\pi(\langle b, \mathfrak{B}\rangle)(p) = \pi(\langle b, \mathfrak{C}\rangle)(p)$, therefore $(\mathbb{M}, \langle b, \mathfrak{C}\rangle) \models \neg p$.

Third, consider the case where $\phi \wedge \psi \in \mathfrak{C}$. I have then that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \phi$ and $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \psi$. However, by the inductive step I already have shown that $(\mathbb{M}, \langle b, \mathfrak{C}\rangle) \models \phi$ and $(\mathbb{M}, \langle b, \mathfrak{C}\rangle) \models \psi$; therefore $(\mathbb{M}, \langle b, \mathfrak{C}\rangle) \models \phi \wedge \psi$.

Finally, consider the case of $\neg\phi$. If ϕ is of the form $\neg\psi$ then $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \neg\neg\psi$ if and only if $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \psi$, and by the inductive step $(\mathbb{M}, \langle b, \mathfrak{C}\rangle) \models \psi$ and finally $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \neg\neg\psi$. If ϕ is of the form $\neg(\psi \wedge \chi)$, then either $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \neg\psi$ or $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \neg\phi$. By the inductive step I have that either $(\mathbb{M}, \langle b, \mathfrak{C}\rangle) \models \neg\psi$ or $(\mathbb{M}, \langle b, \mathfrak{C}\rangle) \models \neg\phi$, and therefore $(\mathbb{M}, \langle b, \mathfrak{C}\rangle) \models \neg(\phi \wedge \psi)$ $\square$

Theorem 4.6.3 $\pm$ *Theorems*

(*i*) $\models \boxplus\phi \rightarrow \boxplus\boxplus\,\phi$

(*ii*) $\models \boxplus\phi \rightarrow \boxminus\boxplus\,\phi$

(*iii*) $\models \boxminus\phi \rightarrow \boxminus\boxminus\,\phi$

(*iv*) $\models \boxplus\boxminus\bot$

Pf.

(i) Consider any worlds $\langle b, \mathfrak{B}\rangle$, $\langle b, \mathfrak{C}\rangle$, $\langle b, \mathfrak{D}\rangle$ such that $R_+(\langle b, \mathfrak{B}\rangle, \langle b, \mathfrak{C}\rangle)$ and $R_+(\langle b, \mathfrak{C}\rangle, \langle b, \mathfrak{D}\rangle)$. It suffices to show $R_+(\langle b, \mathfrak{B}\rangle, \langle b, \mathfrak{D}\rangle)$, for if $R_+(\langle b, \mathfrak{B}\rangle, \langle b, \mathfrak{D}\rangle)$ and $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \boxplus\phi$, then $(\mathbb{M}, \langle b, \mathfrak{D}\rangle) \models \phi$, and therefore $(\mathbb{M}, \langle b, \mathfrak{C}\rangle) \models \boxplus\phi$, and therefore $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \boxplus\boxplus\,\phi$. To show $R_+(\langle b, \mathfrak{B}\rangle, \langle b, \mathfrak{D}\rangle)$, as $\mathfrak{D} \subseteq \mathfrak{C} \subseteq \mathfrak{B}$, it suffices to show $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mathfrak{D}$. However, because $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mathfrak{C}$ by assumption, and $\mathfrak{D} \subseteq \mathfrak{C}$, then $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mathfrak{D}$, which suffices to prove the theorem.

(ii) It suffices to show if $R_-(\langle b, \mathfrak{B}\rangle, \langle b, \mathfrak{C}\rangle)$ and $R_+(\langle b, \mathfrak{C}\rangle, \langle b, \mathfrak{D}\rangle)$ then $R_+(\langle b, \mathfrak{B}\rangle, \langle b, \mathfrak{D}\rangle)$. So I must show if $(\mathbb{M}\langle b, \mathfrak{C}\rangle) \models \mathfrak{D}$ then $(\mathbb{M}\langle b, \mathfrak{B}\rangle) \models \mathfrak{D}$. The proof is complete by the $\pm$ lemma.

(iii) It suffices to show if $R_-(\langle b, \mathfrak{B}\rangle, \langle b, \mathfrak{C}\rangle)$ and $R_-(\langle b, \mathfrak{C}\rangle, \langle b, \mathfrak{D}\rangle)$ then $R_-(\langle b, \mathfrak{B}\rangle, \langle b, \mathfrak{D}\rangle)$. So I must show if $(\mathbb{M}\langle b, \mathfrak{C}\rangle) \not\models \mathfrak{D}$ then $(\mathbb{M}\langle b, \mathfrak{B}\rangle) \not\models \mathfrak{D}$. The proof follows again the $\pm$ lemma.

(iv) It suffices to show if $R_+(\langle b, \mathfrak{B}\rangle, \langle b, \mathfrak{C}\rangle)$ then $R_+(\langle b, \mathfrak{C}\rangle, \langle b, \mathfrak{C}\rangle)$ that is, to show if $(\mathbb{M}\langle b, \mathfrak{B}\rangle) \models \mathfrak{C}$ then $(\mathbb{M}\langle b, \mathfrak{C}\rangle) \models \mathfrak{C}$. Once again, the $\pm$ lemma proves the result.

$\square$

The philosophical interpretation of the $\pm$ theorem is as follows:

(*i*) $\models \boxplus\phi \rightarrow \boxplus\boxplus\,\phi$ states that if the agent with a belief basis $\mathfrak{B}$ has a subset $\mathfrak{C} \subseteq \mathfrak{B}$ which is grounded, and $\mathfrak{D} \subseteq \mathfrak{C}$ is grounded when one considers $\mathfrak{C}$ as a belief basis, then $\mathfrak{D}$ is a grounded subset of $\mathfrak{B}$.

(*ii*) $\models \boxplus\phi \rightarrow \boxminus\boxplus\,\phi$ states that if the agent with a belief basis $\mathfrak{B}$ has a subset $\mathfrak{C} \subseteq \mathfrak{B}$ which is specious, and $\mathfrak{D} \subseteq \mathfrak{C}$ is grounded when one considers $\mathfrak{C}$ as a belief basis, then $\mathfrak{D}$ is a grounded subset of $\mathfrak{B}$.

(*iii*) $\models \boxminus\phi \rightarrow \boxminus\boxminus\,\phi$ states that if the agent with a belief basis $\mathfrak{B}$ has a subset $\mathfrak{C} \subseteq \mathfrak{B}$ which is specious, and $\mathfrak{D} \subseteq \mathfrak{C}$ is specious when one considers $\mathfrak{C}$ as a belief basis, then $\mathfrak{D}$ is a specious subset of $\mathfrak{B}$.

(*iv*) $\models \boxplus\boxminus\bot$ states that if the agent with a belief basis $\mathfrak{B}$ has a subset $\mathfrak{C} \subseteq \mathfrak{B}$ which is grounded, then there are no specious subsets when one considers $\mathfrak{C}$ as a belief basis.

THEOREM 4.6.4 *Speciousness Theorem*

$$\models \diamondminus\phi \rightarrow (\boxminus\psi \rightarrow \psi)$$

Pf. Assuming $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondminus\phi$, it suffices to show that if $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \boxminus\psi$, then $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \psi$, so suppose that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \boxminus\psi$. Then, for all worlds $\langle b, \mathfrak{C}\rangle$ such that $R_{-}(\langle b, \mathfrak{B}\rangle, \langle b, \mathfrak{C}\rangle)$, then $(\mathbb{M}, \langle b, \mathfrak{C}\rangle) \models \psi$.

If $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondminus\phi$, then there is some $\mathfrak{C} \subseteq \mathfrak{B}$ such that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \not\models \mathfrak{C}$ and $(\mathbb{M}, \langle b, \mathfrak{C}\rangle) \models \phi$. This means, however, that for some $\chi \in \mathfrak{C}$ that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \not\models \chi$. However, if $\chi \in \mathfrak{C}$ then $\chi \in \mathfrak{B}$, which means that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \not\models \mathfrak{B}$. That just means that $R_{-}(\langle b, \mathfrak{B}\rangle, \langle b, \mathfrak{B}\rangle)$. Thus I have that if $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \boxminus\psi$, then $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \psi$. □

Corollary 4.6.5

$$\models (B\phi \wedge \diamondminus\psi) \to \diamondminus B\phi$$

Pf. Suppose that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models B\phi \wedge \diamondminus\psi$, it suffices to show that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondminus B\phi$. If $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models B\phi \wedge \diamondminus\psi$ then $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondminus\psi$, so $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \boxminus\neg B\phi \to \neg B\phi$ by the speciousness theorem. This is equivalent to stating that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models B\phi \to \diamondminus B\phi$. However, because $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models B\phi \wedge \diamondminus\psi$, that means that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models B\phi$. Therefore $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondminus B\phi$, which was what I was out to prove. □

To give a philosophical reading of this corollary, it states that if the agent believes ϕ, and they have a false belief (that is, they have a subset of their belief basis which is specious) then they have a possibly specious belief that ϕ. In other words, this gives a sufficient condition for determining when the agent has a possibly specious belief of a propositions ϕ.

The speciousness theorem, while deeper, is much more difficult to characterize philosophically. To employ modal logic intuition, it states that whenever the agent has a a false belief then R_{-} is reflexive. This theorem may be rewritten to say $(\boxminus\psi \to \diamondminus\psi) \to (\boxminus\phi \to \phi)$; I suspect that an even more elegant way exists of expressing it but I am not certain.

Corollary 4.6.6

(*i*) $\models (\diamondminus B\phi \wedge \diamondminus B(\phi \to \psi)) \to \diamondminus B\psi$

(*ii*) $\models (\diamondplus B\phi \wedge \diamondminus B(\phi \to \psi)) \to \diamondminus B\psi$

(*iii*) $\models (\diamondminus B\phi \wedge \diamondplus B(\phi \to \psi)) \to \diamondminus B\psi$

Pf. I will endeavor to prove the first case without loss of generality. It suffices to show if $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondminus B\phi \wedge \diamondminus B(\phi \to \psi)$ then $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondminus B\psi$. If $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondminus B\phi \wedge \diamondminus B(\phi \to \psi)$ then $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models B\phi \wedge B(\phi \to \psi)$ by the belief theorem, and therefore $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models B\psi$. However, because I have $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondminus B\phi \wedge \diamondminus B(\phi \to \psi)$ then I have that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondminus B\phi$, and therefore $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models B\psi \wedge \diamondminus B\phi$. By the corollary of the speciousness theorem, I therefore have that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondminus B\psi$. □

I feel the philosophical motivation for these statements is simpler than the speciousness theorem. These statements can be thought of as *possibly specious belief composition rules*. The first statement,

$\models (\diamondminus B\phi \wedge \diamondminus B(\phi \to \psi)) \to \diamondminus B\psi$, says if the agent can compose a specious argument for ϕ and a specious argument for $\phi \to \psi$, then the agent can compose a specious argument for ψ. The second and third statements correspond to the intuition that even if the agent can make a sound argument for ϕ or $\phi \to \psi$, if they used specious reasoning elsewhere then they can combine these results to arrive at other specious reasoning.

4.6.2 Grounding Theorem

In this section I give a result I call the grounding theorem. It is similar in character to the speciousness theorem. I will give a preliminary definition, state and prove the theorem, and finally provide philosophical commentary.

To give the grounding theorem, I first need to state a definition:

Definition 4.6.1 A model $\mathbb{M}$ is said to be *closed under union of grounds* if and only if whenever $\langle b, \mathfrak{B}\rangle, \langle b, \mathfrak{C}\rangle \in \Omega$, $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models\!\models \mathfrak{B}$ and $(\mathbb{M}, \langle b, \mathfrak{C}\rangle) \models\!\models \mathfrak{C}$, then there exists a world $\langle b, \mathfrak{B} \cup \mathfrak{C}\rangle \in \Omega$

Theorem 4.6.7 If $\mathbb{M}$ is closed under union of grounds, then

$$\mathbb{M} \models (\diamondplus B\phi \wedge \diamondplus B(\phi \to \psi)) \to \diamondplus B\psi$$

Pf. To show this theorem, suppose that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondplus B\phi \wedge \diamondplus B(\phi \to \psi)$. It suffices to show that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondplus B\psi$.

The hypothesis $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondplus B\phi \wedge \diamondplus B(\phi \to \psi)$ says that there are two worlds $\langle b, \mathfrak{C}\rangle, \langle b, \mathfrak{D}\rangle$ such that $R_+(\langle b, \mathfrak{B}\rangle, \langle b, \mathfrak{C}\rangle)$, $R_+(\langle b, \mathfrak{B}\rangle, \langle b, \mathfrak{D}\rangle)$, and:

(i) If $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models \mathfrak{C}$ then $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models \phi$

(ii) If $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models \mathfrak{D}$ then $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models \phi \to \psi$

I know from closure under grounds that there is a world $\langle b, \mathfrak{C} \cup \mathfrak{D}\rangle \in \Omega$. I hold that not only is it true (1) that $(\mathbb{M}, \langle b, \mathfrak{C} \cup \mathfrak{D}\rangle) \models B\psi$, but also that (2) $R_+(\langle b, \mathfrak{B}\rangle, \langle b, \mathfrak{C} \cup \mathfrak{D}\rangle)$, which suffice to show that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondplus B\psi$.

To show both (1) and (2), it is helpful to see that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mathfrak{X}$ and $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mathfrak{Y}$ if and only if $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mathfrak{X} \cup \mathfrak{Y}$.

First, I will endeavor to show if $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mathfrak{X}$ and $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mathfrak{Y}$ then $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mathfrak{X} \cup \mathfrak{Y}$. This is because for all $\phi \in \mathfrak{X} \cup \mathfrak{Y}$, either $\phi \in \mathfrak{X}$ or $\phi \in \mathfrak{Y}$, and because $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mathfrak{X}$ and $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mathfrak{Y}$ thus means $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \phi$ in either case.

Second, I will endeavor to show if $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mathfrak{X} \cup \mathfrak{Y}$, then $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mathfrak{X}$ and $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mathfrak{Y}$. If $\phi \in \mathfrak{X}$, then $\phi \in \mathfrak{X} \cup \mathfrak{Y}$, and thus $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \phi$ by hypothesis. Therefore $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mathfrak{X}$; a similar argument works to show $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mathfrak{Y}$.

To show (1), it suffices to show that if $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models \mathfrak{C} \cup \mathfrak{D}$, then $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models \phi$ and $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models \phi \to \psi$. Suppose that $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models\!\models \mathfrak{C} \cup \mathfrak{D}$, then $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models\!\models \mathfrak{C}$ and therefore $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models\!\models$

ϕ by hypothesis. Similarly, if $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models\models \mathfrak{C} \cup \mathfrak{D}$ then $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models\models \mathfrak{D}$ and therefore $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models\models \phi \to \psi$. But if $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models \phi$ and $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models \phi \to \psi$, then $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models \psi$.

To show (2), just note that by hypothesis $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mathfrak{C}$ and $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mathfrak{D}$, and therefore $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mathfrak{C} \cup \mathfrak{D}$. □

A word should be provided for this theorem. This says that, when $\mathbb{M}$ has specific properties, if the agent can compose a sound argument for ϕ and compose another sound argument for $\phi \to \psi$, then the agent can compose yet another sound argument for ψ. Recall that it was not necessary to stipulate anything beyond the conventional semantics of FL to arrive at this result for $\mathbb{M}$. This is because if the agent has a false belief, then the sum of her beliefs is specious, and so any reasoning from the entirety of their beliefs may be possibly specious. The agent can, under these semantics, always reason from all of her experience, so no special stipulations are necessary. However, without specious properties, she might not be able to chain sound arguments together.

4.6.3 Deductive Closure Theorem

In this section I provide the a consequence I call the deductive closure theorem. This section draws upon the definition of a belief base, which was provided in §2.1.2. It states that, under sufficient conditions, the deductive closure of an agent's belief basis generates her belief base. To illustrate this theorem, I will first give a definition of a belief base at a particular world, and then state the theorem. I will briefly reflect philosophically on this result at the end of this section.

Recall the definition of *deductive closure* in Def. 2.2.2 in §2.2.2. At a particular world, I may define an agent's *belief base* as thus:

Definition 4.6.2 Let $\langle b, \mathfrak{B}\rangle$ be some world in $\mathbb{M}$. Then an agent's belief base at $\langle b, \mathfrak{B}\rangle$ is defined to be:

$$\mathscr{B}_{\langle b,\mathfrak{B}\rangle} = \{\phi \mid (\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models B\phi\}$$

Theorem 4.6.8 For any world $\langle b, \mathfrak{B}\rangle$ in any model $\mathbb{M}$, $\overline{\mathfrak{B}} \subseteq \mathscr{B}_{\langle b,\mathfrak{B}\rangle}$. If $\mathfrak{B}$ is finite, then $\overline{\mathfrak{B}} = \mathscr{B}_{\langle b,\mathfrak{B}\rangle}$.

Pf. I first prove $\overline{\mathfrak{B}} \subseteq \mathscr{B}_{\langle b,\mathfrak{B}\rangle}$. It suffices to show that $\mathfrak{B}$ satisfies Defs. 2.2.2(i), 2.2.2(ii), and 2.2.2(iii), but not necessarily the minimality condition.

Def. 2.2.2(*i*) Let $\phi \in \mathfrak{B}$. By definition, $R_B(\langle b, \mathfrak{B}\rangle, \langle x, \mathfrak{X}\rangle)$ if and only if $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models \mathfrak{B}$. Therefore, if $R_B(\langle b, \mathfrak{B}\rangle, \langle x, \mathfrak{X}\rangle)$ then $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models \phi$. Thus, $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models B\phi$, and $\phi \in \mathscr{B}_{\langle b,\mathfrak{B}\rangle}$.

Def. 2.2.2(*ii*) Recall that $\mathscr{T}_{\mathbb{M}} = \{\phi \mid \mathbb{M} \models \phi\}$. From Prop. 4.6.1($iii$), I have that if $\mathbb{M} \models \phi$, then $\mathbb{M} \models B\phi$. Therefore, $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models B\phi$, and therefore $\phi \in \mathscr{B}_{\langle b,\mathfrak{B}\rangle}$. Thus, $\mathscr{T}_{\mathbb{M}} \subseteq \mathscr{B}_{\langle b,\mathfrak{B}\rangle}$

Def. 2.2.2(*iii*) The proof proceeds inductively; the base case is just the case that $\phi \in \mathfrak{B}$ or $\phi \in \mathscr{T}_{\mathbb{M}}$, which I have already demonstrated.

Suppose I have shown for $\phi \in \overline{\mathfrak{B}}$, $\phi \to \psi \in \overline{\mathfrak{B}}$ that $\phi \in \mathscr{B}_{\langle b,\mathfrak{B}\rangle}$ and $\phi \to \psi \in \mathscr{B}_{\langle b,\mathfrak{B}\rangle}$. I want to show $\psi \in \mathscr{B}_{\langle b,\mathfrak{B}\rangle}$. From $\phi \in \mathscr{B}_{\langle b,\mathfrak{B}\rangle}$ and $\phi \to \psi \in \mathscr{B}_{\langle b,\mathfrak{B}\rangle}$ I may conclude that

$(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models B\phi$ and $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models B(\phi \to \psi)$. Therefore $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models B\phi \wedge B(\phi \to \psi)$. However, I know that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models (B\phi \wedge B(\phi \to \psi)) \to B\psi$. Therefore, I can conclude that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models B\psi$, which means that $\psi \in \mathscr{B}_{\langle b, \mathfrak{B}\rangle}$.

Next, I prove if $\mathfrak{B}$ is finite, then $\overline{\mathfrak{B}} = \mathscr{B}_{\langle b, \mathfrak{B}\rangle}$. It suffices to show $\overline{\mathfrak{B}} \supseteq \mathscr{B}_{\langle b, \mathfrak{B}\rangle}$. Let $\mathfrak{B} = \{\chi_1, \chi_2, \ldots, \chi_n\}$; let μ denote $\chi_1 \wedge \chi_2 \wedge \ldots \wedge \chi_n$. Notice that the following is a tautology:

$$\chi_1 \to (\chi_2 \to (\ldots(\chi_n \to (\mu))\ldots))$$

Therefore, by Def. 2.2.2(iii), $\mu \in \overline{\mathfrak{B}}$ and $\mu \in \mathscr{B}_{\langle b, \mathfrak{B}\rangle}$ by above. Furthermore, for all worlds $\langle x, \mathfrak{X}\rangle$, $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models \mathfrak{B}$ if and only if $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models \mu$.

I argue that if $\psi \in \mathscr{B}_{\langle b, \mathfrak{B}\rangle}$, then $\mathbb{M} \models \mu \to \psi$. If $\psi \in \mathscr{B}_{\langle b, \mathfrak{B}\rangle}$, then $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models B\psi$, which means that for all $\langle x, \mathfrak{X}\rangle$ such that $R_B(\langle b, \mathfrak{B}\rangle, \langle x, \mathfrak{X}\rangle)$ then $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models \psi$. Thus, if $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models \mathfrak{B}$, then $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models \psi$. However, this means that if $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models \mu$, then $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models \psi$, which just means that $\mathbb{M} \models \mu \to \psi$.

As $\mathscr{T}_{\mathbb{M}} \subseteq \overline{\mathfrak{B}}$, then $\mu \to \psi \in \overline{\mathfrak{B}}$ as well as $\mu \in \overline{\mathfrak{B}}$. Therefore, $\psi \in \overline{\mathfrak{B}}$, which means that $\mathscr{B}_{\langle b, \mathfrak{B}\rangle} \subseteq \overline{\mathfrak{B}}$. □

I feel this theorem deserves a philosophical explanation. It states that the deductive closure of an agent's belief basis will always be contained in the agent's belief base. It goes on to state that if the agent has only a finite number of expressions in their belief basis, then their belief basis generates their belief base. Note that this is not a *necessary* condition. To see this, consider the following belief basis:

$$\mathfrak{B} = \{p, (p \to p) \to p, (p \to p) \to ((p \to p) \to p), \\ (p \to p) \to ((p \to p) \to ((p \to p) \to p)), \ldots\}$$

...with a little work it can be demonstrated that $\overline{\mathfrak{B}} = \overline{\{p\}}$, and furthermore that $\mathscr{B}_{\langle b, \mathfrak{B}\rangle} = \mathscr{B}_{\langle b, \{p\}\rangle}$. To give jointly necessary and sufficient conditions for $\overline{\mathfrak{B}} = \mathscr{B}_{\langle b, \mathfrak{B}\rangle}$, I hold that $\mathscr{B}_{\langle b, \mathfrak{B}\rangle}$ needs to be *compact* with respect to $\mathfrak{B} \cup \mathscr{T}_{\mathbb{M}}$. An reader interested in this subject may care to look at an elementary text in model theory for further information (for instance, see Enderton, 1972).

4.6.4 Cartesian Theorem

In this section I present the cartesian theorem. This theorem owes its inspiration to René Descartes (1996) in *Meditations II*. I will present the theorem, present a relevant definition and mention philosophical interpretation.

THEOREM 4.6.9 Suppose that $\langle b, \varnothing\rangle \in \Omega$, then $\mathbb{M} \models \phi$ if and only if $(\mathbb{M}, \langle b, \varnothing\rangle) \models B\phi$

Pf. Suppose that $\mathbb{M} \models \phi$, then for any $(\mathbb{M}, \langle c, \mathfrak{C}\rangle) \models \varnothing$ I have that $(\mathbb{M}, \langle c, \mathfrak{C}\rangle) \models \phi$, which means that $(\mathbb{M}, \langle b, \varnothing\rangle) \models B\phi$. Conversely, if $(\mathbb{M}, \langle b, \varnothing\rangle) \models B\phi$ then for any world such that $(\mathbb{M}, \langle c, \mathfrak{C}\rangle) \models \varnothing$, then $(\mathbb{M}, \langle c, \mathfrak{C}\rangle) \models \phi$. However, I have that $\models \varnothing$ holds vacuously. Therefore, this means that $\mathbb{M} \models \phi$. □

I now turn to defining what it means for a model to be *Cartesian*:

DEFINITION 4.6.3 $\mathbb{M}$ is *Cartesian* if and only if for all $\langle b, \mathfrak{B}\rangle \in \Omega$, $\langle b, \varnothing\rangle \in \Omega$

I now turn to motivating philosophically the idea of what it means for a model to be Cartesian. As stated, the observation that the empty set is a grounded subset of one's experiential beliefs is due to René Descartes (1996) in *Meditations II*:

> The Meditation of yesterday has filled my mind with so many doubts, that it is no longer in my power to forget them. Nor do I see, meanwhile, any principle on which they can be resolved; and, just as if I had fallen all of a sudden into very deep water, I am so greatly disconcerted as to be unable either to plant my feet firmly on the bottom or sustain myself by swimming on the surface. I will, nevertheless, make an effort, and try anew the same path on which I had entered yesterday, that is, proceed by *casting aside all that admits of the slightest doubt, not less than if I had discovered it to be absolutely false*; and I will continue always in this track until I shall find something that is certain, or at least, if I can do nothing more, until I shall know with certainty that there is nothing certain.

I hold that it is a powerful result when it is possible, in all cases, to cut away all experiential beliefs the agent has and deal with nothing more than their background knowledge. For one thing, background knowledge is easy to characterize modally; indeed, this is what traditional epistemic logic largely studies. Suppose I had a new accessibility relation $R_* = \Omega \times \Omega$ which governs $\boxast$, then it may be seen that $\boxast$ is symmetric, reflexive and transitive, and furthermore $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \boxast\phi$ if and only if $\phi \in \mathscr{T}_{\mathbb{M}}$. Essentially, $\boxast$ means something like *metaphysical necessity* for the system the agent is reasoning in. It also be seen, if $\mathbb{M}$ is Cartesian, that $\boxast\phi \leftrightarrow \boxplus B\phi$. Thus, $\boxplus B$ may be thought of as basically operating together in an **S5** system. I feel that thinking about an agent when they are reasoning from nothing is is like thinking about what the old philosopher was trying to accomplish in his *Meditations*: when an agent discards all experiential beliefs then she can, in a sense, reason purely based on only what is necessary. It may be remarked, however, that I cannot show that an agent may reason about its own existence in my logic, nor about God, to the best of my knowledge.

4.6.5 Infinite Falsity Lemma

In this section I present a result I call the *infinite falsity lemma*[7]. It amounts to an analysis of a statement which entails that "the agent believes an infinite number of false things." I will first present a definition which will be helpful for the proof of my result, then state and prove the theorem. I will then demonstrate a counter-example to the converse of this theorem, and finally remark philosophically on the nature of this result.

For this lemma, it will be useful to have a definition:

[7] I would like to thank Dan Linford for helping me to arrive at this result, through conversations I had with him during the writing of this thesis.

DEFINITION 4.6.4 $\mathfrak{C}$ is said to be *totally false* in $\langle b, \mathfrak{B}\rangle$ if $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \not\models \phi$ for all $\phi \in \mathfrak{C}$.

LEMMA 4.6.2 If $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \boxminus B\phi \to \phi$ for all $\phi \in \mathscr{L}_{\mathsf{FL}}$, then for every integer m there exists a $\mathfrak{C} \subseteq \mathfrak{B}$ such that $m < |\mathfrak{C}|$ which is totally false in $\langle b, \mathfrak{B}\rangle$.

Pf. Assume $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \boxminus B\phi \to \phi$, and suppose, for purposes of *reductio ad absurdum*, that it is not the case that for every integer m there exists a $\mathfrak{C} \subseteq \mathfrak{B}$ such that $|\mathfrak{C}| > m$ which is totally false in $\langle b, \mathfrak{B}\rangle$. This can be seen to be equivalent to supposing that there is an integer m such that every subset $\mathfrak{X} \subseteq \mathfrak{B}$ which is totally false in $\langle b, \mathfrak{B}\rangle$ contains at most m elements. Let $\mathfrak{M}$ be the union of the totally false subsets; it can be seen that $\mathfrak{M}$ is totally itself so therefore $|\mathfrak{M}| \leq m$, and if $\mathfrak{D}$ is totally false then $|\mathfrak{D}| \leq |\mathfrak{M}|$.

I argue every specious subset $\mathfrak{F}$ of $\mathfrak{B}$ intersects $\mathfrak{M}$. Suppose that there was a specious $\mathfrak{F}$ where $\mathfrak{F} \cap \mathfrak{M} = \varnothing$. Since $\mathfrak{F}$ is specious, $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \not\models \mathfrak{F}$, and therefore there is some $\psi \in \mathfrak{F}$ such that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \not\models \psi$. However, if $\mathfrak{M}$ is totally false, then $\mathfrak{M} \cup \{\psi\}$ is also totally false. However, $|\mathfrak{M} \cup \{\psi\}| = |\mathfrak{M}| + 1$, which contradicts that if $\mathfrak{D}$ is totally false then $|\mathfrak{D}| \leq |\mathfrak{M}|$.

Let $\mathfrak{M} = \{\chi_1, \chi_2, \ldots, \chi_n\}$, and let μ denote $\chi_1 \vee \chi_2 \vee \ldots \vee \chi_n$, the disjunction of all of the elements in $\mathfrak{M}$. This is ensured to be a formula in $\mathscr{L}_{\mathsf{FL}}$ because $|\mathfrak{M}|$ is less than some integer m by assumption and therefore finite. I argue $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \boxminus B\mu$. To show this, it suffice to consider any specious $\mathfrak{F}$ and show $(\mathbb{M}, \langle b, \mathfrak{F}\rangle) \models B\mu$.

Since $\mathfrak{F}$ is specious, $\mathfrak{F} \cap \mathfrak{M} \neq \varnothing$, so take some $\chi_i \in \mathfrak{F} \cap \mathfrak{M}$. Let $\langle g, \mathfrak{G}\rangle$ be a world such that $R_B(\langle b, \mathfrak{F}\rangle, \langle g, \mathfrak{G}\rangle)$, then $(\mathbb{M}, \langle g, \mathfrak{G}\rangle) \models \chi_i$ by definition, and therefore $(\mathbb{M}, \langle g, \mathfrak{G}\rangle) \models \mu$. Therefore, $(\mathbb{M}, \langle b, \mathfrak{F}\rangle) \models B\mu$, which suffices to show that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \boxminus B\mu$.

I assumed that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \boxminus B\phi \to \phi$, so therefore $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \boxminus B\phi \to \phi$ where ϕ was any sentence in $\mathscr{L}_{\mathsf{FL}}$, so therefore (I) $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mu$.

However, note that if $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \not\models \alpha$ and $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \not\models \beta$ then $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \not\models \alpha \vee \beta$. Since $\mathbb{M}$ is totally false, then $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \not\models \chi_i$ for all $1 \leq i \leq n$, then by induction (II) $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \not\models \mu$.

From (I) and (II), I therefore have $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mu$ and $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \not\models \mu$. Therefore, the assumption I took for reductio must be false, which suffices to prove the theorem. □

Next note that the converse of this statement is not true; to see this I provide the following example:

EXAMPLE 4.6.1 Let $\mathbf{P} = \{q\} \cup \{P \mid i \text{ is an integer}\}$. Construct the following modified Kripke structure

$$\mathbb{M} = \langle \Omega, \pi, R_K, R_B, R_+, R_- \rangle$$

Let $S = \{b, c\} \times \mathscr{P}\mathfrak{B}$ where $\mathfrak{B} = \{q, q \wedge p_1, q \wedge p_2, q \wedge p_3, \ldots\}$, $\pi(\langle b, \mathfrak{X}\rangle)(q) = \mathsf{False}$ and $\pi(\langle b, \mathfrak{X}\rangle)(P) = \mathsf{True}$ for all i and for all $\mathfrak{X}$, and $\pi(\langle c, \mathfrak{X}\rangle)(x) = \mathsf{True}$ for all $x \in \mathbf{P}$ and for all $\mathfrak{X}$. Let $R_K = \varnothing$, and $R_B, R_+,$ andR_- be as defined. Clearly $\mathfrak{B}$ is totally false at $\langle b, \mathfrak{B}\rangle$, however $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \boxminus Bq$ and $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \not\models q$.

Finally, I remark on the philosophical nature of this theorem. First of all, finding words in ordinary language, and even modal logic to express this theorem is difficult. It is not clear what it means when $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \boxminus B\phi \to \phi$ is true for all $\phi \in \mathscr{L}_{\mathsf{FL}}$; however I have some theories.

Let $\mathscr{F}_{\langle b,\mathfrak{B}\rangle} = \{\phi \mid (\mathbb{M}, \langle b,\mathfrak{B}\rangle) \models \boxminus B\phi\}$. Using the notation from the previous section, this means that $\mathscr{F}_{\langle b,\mathfrak{B}\rangle} = \bigcap\{\mathscr{B}_{\langle b,\mathfrak{F}\rangle} \mid R_-(\langle b,\mathfrak{B}\rangle, \langle b,\mathfrak{F}\rangle)\}$, namely that $\mathscr{F}_{\langle b,\mathfrak{B}\rangle}$ is the intersection of all of the belief bases generated by specious subsets of $\mathfrak{B}$ at $\langle b,\mathfrak{B}\rangle$ If $(\mathbb{M}, \langle b,\mathfrak{B}\rangle) \models \boxminus B\phi \to \phi$ is true for all $\phi \in \mathscr{L}_{\mathsf{FL}}$, then that means $(\mathbb{M}, \langle b,\mathfrak{B}\rangle) \models \mathscr{F}_{\langle b,\mathfrak{B}\rangle}$. It may be demonstrated that this has to mean $|\{\mathfrak{F} \mid R_-(\langle b,\mathfrak{B}\rangle, \langle b,\mathfrak{F}\rangle)\}|$ must be infinite. What is going on in this theorem is terrifically complicated; it largely amounts from the fact that $(\mathbb{M}, \langle b,\mathfrak{B}\rangle) \models \boxminus B\phi \to \phi$ states that the agent has committed a feat similar to deriving wine from vinegar. Finding jointly necessary and sufficient conditions for $(\mathbb{M}, \langle b,\mathfrak{B}\rangle) \models \boxminus B\phi \to \phi$ to obtain poses an exciting frontier of exploration for FL.

4.7 Possible Axioms

I turn now from the core semantics of FL, and consider possible extensions, which reflect philosophical ideas I have already put forward.

4.7.1 The (FL1) Axiom

Proposition 4.7.1 $\mathbb{M}$ satisfies axiom (**FL1**) if and only if

$$\mathbb{M} \models K\phi \to \diamondplus B\phi$$

I hold that this axiom encompasses the notion that beliefs that depend *essentially* on falsity, according to Def. **??** in my preliminary philosophical research, are not knowledge. If a belief depends essentially on falsity then it seems intuitive to write this as $B\phi \wedge \neg\diamondplus B\phi$. From the belief theorem, if an ideal agent believes ϕ at $\langle s,\mathfrak{B}\rangle$, then they either believe it on the basis of some grounded subset of $\mathfrak{B}$ or they believe it on some specious subset of $\mathfrak{B}$. If there is no grounded subset of $\mathfrak{B}$ which serves as a belief basis, then it can only have followed from specious subsets of $\mathfrak{B}$, which I ruled out philosophically in §**??**.

Some immediate consequences follow if $\mathbb{M}$ satisfies (**FL1**):

$$\mathbb{M} \models K\phi \to \phi \tag{4.7.1}$$

...which follows from the justification theorem. This reflects that if one knows a proposition, then it is fact true, which I stipulated in Def. 2.1.1(c), and reflects (**EL2**) from epistemic logic. Furthermore:

$$\mathbb{M} \models K\phi \to B\phi \tag{4.7.2}$$

...which follows from the belief theorem. This was another philosophical proposition I committed myself to, namely Prop. 2.2.1, which reflects axiom (**EL5**).

4.7.2 The (FL2) Axiom

Proposition 4.7.2 $\mathbb{M}$ satisfies axiom (**FL2**) if and only if

$$\mathbb{M} \models (\diamondplus B\phi \wedge \boxminus(B\phi \to \diamondplus B\phi)) \to K\phi$$

This axiom, which I am tempted to call axiom **FL**, states that if one's belief of some proposition ϕ is *fully grounded*, as per Def. 2.2.7, then one has knowledge of ϕ. This is just . I consider this to be the most original thing FL logic has to say - if ϕ is never a false lemma, and the agent believes it, then the agent knows ϕ, just as I philosophized in Prop. 2.2.4.

4.7.3 The (FL3) Axiom

PROPOSITION 4.7.3 $\mathbb{M}$ satisfies axiom (**FL3**) if and only if

$$\mathbb{M} \models \neg B\bot$$

This axiom holds that the ideal agent under investigation holds no contradictory beliefs, reflecting 2.1.2(c). Note that it reflects the axiom (**EL5**) from the epistemic logic literature.

This axiom also poses something fascinating, for one can think of a world $\langle b, \mathfrak{B}\rangle$ as being a model of a belief basis $\mathfrak{C}$ in some other world $\langle c, \mathfrak{C}\rangle$ that has doxastic access to $\langle b, \mathfrak{B}\rangle$. Thus, this axiom says that every $\mathfrak{B}$ such that there is a $\langle b, \mathfrak{B}\rangle \in \Omega$ has a model in $\mathbb{M}$. Despite the circularity of this axiom, I hold that, as opposed to the other two axioms, classification of models satisfying this axiom is possible in the semantics I have outlined.

4.8 Negative Doxastic Introspection

4.8.1 Definition

I state now *negative introspection*, with the intention of giving an argument against it:

PROPOSITION 4.8.1 $\mathbb{M}$ satisfies axiom (**FL4**) (*negative doxastic introspection*) if and only if

$$\mathbb{M} \models \neg B\phi \to B\neg B\phi$$

$\mathbb{M}$ satisfies axiom (**FL5**) if and only if

$$\mathbb{M} \models \neg B\phi \to K\neg B\phi$$

Note that (**FL4**) and (**FL5**) follow if R_B and R_K are non-increasing. Furthermore, if $\mathbb{M} \models K\phi \to B\phi$, such as when $\mathbb{M}$ satisfies 4.7.1, then if (**FL5**) holds then (**FL4**) holds. Because of this, I will focus my analysis on (**FL4**).

4.8.2 Negative Doxastic Introspection Theorem

THEOREM 4.8.1 *The Negative Introspection Theorem*

If $\mathbb{M}$ is *negatively introspective*, that is, (**FL4**) holds, then

$$\mathbb{M} \models B \boxminus \bot$$

Pf. The result follows immediately from the following lemma. $\square$

LEMMA 4.8.1 If $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \neg B\phi \to B\neg B\phi$ for every $\phi \in \mathscr{L}_{\mathsf{FL}}$, then $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models B \boxminus \bot$.

Pf. First observe that if $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models \mathfrak{X}$, then there is no world $\langle y, \mathfrak{Y}\rangle$ such that $R_{-}(\langle x, \mathfrak{X}\rangle, \langle y, \mathfrak{Y}\rangle)$ by definition, which implies that $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models \boxminus\phi$ for all sentences ϕ, including $\bot$. Likewise, if $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models \boxminus\bot$, then there cannot be a world such that $R_{-}(\langle x, \mathfrak{X}\rangle, \langle y, \mathfrak{Y}\rangle)$, as that would imply that some world has the property $(\mathbb{M}, \langle y, \mathfrak{Y}\rangle) \models \bot$, which cannot happen. Therefore, $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models \mathfrak{X}$ if and only if $(\mathbb{M}, \langle x, \mathfrak{X}\rangle) \models \boxminus\bot$.

It suffices to show for any world $\langle b, \mathfrak{B}\rangle$ in $\mathbb{M}$, then $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models B \boxminus \bot$. To show this, by above it suffices to show that if $(\mathbb{M}, \langle c, \mathfrak{C}\rangle) \models \mathfrak{B}$ then $(\mathbb{M}, \langle c, \mathfrak{C}\rangle) \models \mathfrak{C}$.

Suppose, for *reductio ad absurdum*, that $(\mathbb{M}, \langle c, \mathfrak{C}\rangle) \not\models \mathfrak{C}$; this holds if and only if for some $\phi \in \mathfrak{C}$ then $(\mathbb{M}, \langle c, \mathfrak{C}\rangle) \not\models \phi$. Note that because $\phi \in \mathfrak{C}$, then (I) $(\mathbb{M}, \langle c, \mathfrak{C}\rangle) \models B\phi$.

However, $(\mathbb{M}, \langle c, \mathfrak{C}\rangle) \not\models \phi$ is true if and only if $(\mathbb{M}, \langle c, \mathfrak{C}\rangle) \models \neg\phi$. Recall that $(\mathbb{M}, \langle c, \mathfrak{C}\rangle) \models \mathfrak{B}$, therefore $R_B(\langle b, \mathfrak{B}\rangle, \langle c, \mathfrak{C}\rangle)$. Thus I arrive at the result that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \neg B\phi$; for if $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models B\phi$, then $(\mathbb{M}, \langle c, \mathfrak{C}\rangle) \models \phi$, which contradicts my supposition.

As I assumed that $\mathbb{M} \models \neg B\phi \to B\neg B\phi$, then I know from $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \neg B\phi$ that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models B\neg B\phi$. As $R_B(\langle b, \mathfrak{B}\rangle, \langle c, \mathfrak{C}\rangle)$, so therefore (II) $(\mathbb{M}, \langle c, \mathfrak{C}\rangle) \models \neg B\phi$.

From (I) and (II), I therefore have $(\mathbb{M}, \langle c, \mathfrak{C}\rangle) \models B\phi$ and $(\mathbb{M}, \langle c, \mathfrak{C}\rangle) \models \neg B\phi$. Therefore, the assumption I took for reductio, $(\mathbb{M}, \langle c, \mathfrak{C}\rangle) \not\models \mathfrak{C}$, must be false, which means that $(\mathbb{M}, \langle c, \mathfrak{C}\rangle) \models \mathfrak{C}$, which suffices to prove the theorem. $\square$

THEOREM 4.8.2 *Polarization Theorem*
If $\mathbb{M}$ satisfies (**FL4**), then

$$\mathbb{M} \models \boxplus\bot \vee \boxminus\bot$$

Pf. Consider any world $\langle b, \mathfrak{B}\rangle$. It suffices to show if there is a $\mathfrak{C}$ such that $R_{+}(\langle b, \mathfrak{B}\rangle, \langle c, \mathfrak{C}\rangle)$, then $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \boxminus\bot$. If there is such a $\mathfrak{C}$ then $(\mathbb{M}, \langle b, \mathfrak{C}\rangle) \models B \boxminus \bot$ by the negative introspection theorem. Then $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondplus B \boxminus \bot$, and by the justification theorem, $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \boxminus\bot$. $\square$

THEOREM 4.8.3 *Knowledge Collapse Theorem*
Suppose $\mathbb{M}$ satisfies (**FL4**) and either let $\mathbb{M}$ satisfy (**FL1**), then

$$\mathbb{M} \models \boxminus\bot$$

and

$$\mathbb{M} \models B\phi \to \diamondplus B\phi$$

If $\mathbb{M}$ satisfies (**FL4**), (**FL1**), and (**FL2**) then

$$\mathbb{M} \models B\phi \leftrightarrow K\phi$$

Pf.

$\mathbb{M} \models \boxminus\bot$ — I take as given $\models \phi \to \phi$, so $\mathbb{M} \models K(\phi \to \phi)$. Therefore, by (**FL1**), then $\mathbb{M} \models \diamondplus B(\phi \to \phi)$. Therefore $\mathbb{M} \models \boxminus\bot$ by the polarization theorem.

$\mathbb{M} \models B\phi \to \diamondplus B\phi$ — By above, I have that $\mathbb{M} \models \boxminus\neg B\phi$ for all $\phi \in \mathscr{L}_{\mathsf{FL}}$. However, then $\mathbb{M} \models \neg\diamondminus B\phi$ for any ϕ.

Take any world $\langle b, \mathfrak{B}\rangle$, and suppose $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models B\phi$. By the belief theorem I have that $\mathbb{M} \models \diamondplus B\phi \vee \diamondminus B\phi$ from above. However, as I have that $\mathbb{M} \not\models \diamondminus B\phi$, then $\mathbb{M} \models \diamondplus B\phi$. Therefore, $\mathbb{M} \models B\phi \to \phi$.

$\mathbb{M} \models B\phi \leftrightarrow K\phi$ — Now assume (**FL2**) in addition to (**FL4**) and (**FL1**). From (**FL1**) and the belief theorem I already have that $\mathbb{M} \models K\phi \to B\phi$. All that is left is to show that $\mathbb{M} \models B\phi \to K\phi$.

Again, Take any world $\langle b, \mathfrak{B}\rangle$, and suppose $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models B\phi$. By the above, I have that $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondplus B\phi$. From the first thing I proved, I have that $\mathbb{M} \models \boxminus(\phi \to \diamondplus B\phi)$ vacuously. Thus, $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models (\diamondplus B\phi \wedge \boxminus(\phi \to \diamondplus B\phi)) \to K\phi$.

$\square$

4.8.3 Philosophical Interpretation

In this section I provide my philosophical interpretation of the preceding results.

I could have opted to reason about a Cartesian model instead of a model which obeys (**FL1**). This would have allowed me to arrive at the same result. I mainly chose to reason with (**FL1**) out of aesthetic sense.

Negative doxastic introspection is not suitable for a logic for reasoning about grounded belief. Notice that without the addition of any formulae, the polarization theorem says that for every world $\langle b, \mathfrak{B}\rangle$ in a negative introspective has either, for all possible $\mathfrak{C} \subseteq \mathfrak{B}$, then $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \mathfrak{C}$ or for all possible $\mathfrak{C} \subseteq \mathfrak{B}$, then $(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \not\models \mathfrak{C}$. This amounts to saying that, in any doxastic scenario, either the agent can *only* make sound arguments or can *never* make sound arguments. I hold this is inappropriate for a logic regarding belief.

The reason that negative doxastic introspection forces knowledge to collapse into belief is just because the agent has knowledge of background information. The simple addition of an assumption regarding knowledge, namely (**FL1**), to a system with negative doxastic introspection instantly rules out the case in which the agent cannot make sound arguments. In such a system, if there is knowledge, then the agent always has a sound argument for validities, which means apparently that the agent never makes an unsound argument. If the agent can never make an unsound argument, then the addition (**FL2**) says that belief is just the same thing as knowledge.

Thus, the novel bounds I put forward, codified in (**FL1**) and (**FL2**), along with the semantics I derive which were inspired by these bounds, leads me to feel that negative doxastic introspection is inappropriate for my logic.

4.9 Closing Remarks

In this section, I provide my final remarks on the logic I develop.

First, I have a conjecture. Suppose that $\mathbb{M}$ models *positive doxastic introspection*:

$$\mathbb{M} \models B\phi \rightarrow BB\phi$$

... then $\mathbb{M} \models B \boxminus \bot$, just as in the case of negative doxastic introspection. This is regrettable, as this means that a system with positive doxastic introspection undergoes a similar collapse of $B\phi \leftrightarrow K\phi$. I feel it may be remedied by changing the mechanics of R_B as follows:

Proposition 4.9.1 $R_B(\langle b, \mathfrak{B}\rangle, \langle c, \mathfrak{C}\rangle)$ if and only if

(i) $(\mathbb{M}, \langle c, \mathfrak{C}\rangle) \models \mathfrak{B}$

(ii) $\mathfrak{B} \subseteq \mathfrak{C}$

Changing the semantics this way does not affect too much; most if not all of the proofs I present still go through. The only exception I can think of is the deductive closure theorem, where changing the semantics makes it no longer necessarily the case that if $\mathfrak{B}$ is finite then $\overline{\mathfrak{B}} = \mathscr{B}_{\langle b, \mathfrak{B}\rangle}$.

Second, I do not think that negative doxastic introspection may be remedied this way. To remedy negative doxastic introspection, it would appear that I would be forced to remove the necessary condition of $(\mathbb{M}, \langle c, \mathfrak{C}\rangle) \models \mathfrak{B}$ for $R_B(\langle b, \mathfrak{B}\rangle, \langle c, \mathfrak{C}\rangle)$. To remove this condition would defeat the point of my semantics, I feel, so I am not comfortable changing my system this way.

Finally, I imagine that while this logic is rich, it is in a sense introverted and lonely. It allows me to focus utterly on only one agent. I have thought about how to extend this logic to include two agents, but unfortunately I foresee putting statements about other agent's beliefs in a belief basis as leading to paradoxes I struggled with when formulating this logic. All is not lost; for I see that this logic bears great promise for extension.

5 Results and Final Remarks

In this section I give a retrospective on the work that I have provided. I demonstrate how the logic I propose and the results I derive in §4 correspond to the philosophical observations I made in §2. I will then turn to future avenues for my work, including the possibility of soundness and completeness result with applications to philosophy and computer science, and extensions into *higher dimensions*. I will then present the conclusion of my work.

5.1 Philosophical and Logical Equivalence

In this section I systematically demonstrate how all of the philosophical principles and definitions I wrote about in §2 have duals in §4. I demonstrate this wholesale with the following table:

Philosophy		Logic	
		Knowledge	
Def. 2.1.1(a)	*Beliefs are closed under material implication*: if $\phi \in \mathscr{K}$ and $\phi \to \psi \in \mathscr{K}$, then $\psi \in \mathscr{K}$	Prop. 4.6.1(ii)	$\models (K\phi \land K(\phi \to \psi)) \to K\psi$
Def. 2.1.1(b)	*Background knowledge is known*: $\mathscr{T}_{\mathbb{M}} \subseteq \mathscr{K}$	Prop. 4.6.1(iii)	If $\mathbb{M} \models \phi$ then $\mathbb{M} \models K\phi$
Def. 2.1.1(c)	*The truth condition on knowledge*: if $\phi \in \mathscr{K}$, then ϕ must be True in the current state of affairs	Axiom (**FL1**), Eq. (4.7.1)	$\mathbb{M} \models K\phi \to \phi$
		Belief	
Def. 2.1.2(a)	*Beliefs are closed under material implication*: if $\phi \in \mathscr{B}$ and $\phi \to \psi \in \mathscr{B}$, then $\psi \in \mathscr{B}$	Prop. 4.6.1(ii)	$\models (B\phi \land B(\phi \to \psi)) \to B\psi$
Def. 2.1.2(b)	*Background knowledge is believed*: $\mathscr{T}_{\mathbb{M}} \subseteq \mathscr{B}$	Prop. 4.6.1(iii)	If $\mathbb{M} \models \phi$ then $\mathbb{M} \models B\phi$
Def. 2.1.1(c)	*Beliefs are consistent*: if $\neg\phi \in \mathscr{T}_{\mathbb{M}}$, then $\phi \notin \mathscr{B}$	Axiom (**FL3**)	$\mathbb{M} \models \neg B\bot$
		Knowledge & Belief	
Prop. 2.2.1, Eq. (2.2.1)	*Knowledge entails belief*: $\mathscr{K} \subseteq \mathscr{B}$	Axiom (**FL1**), Eq. (4.7.2)	$\mathbb{M} \models K\phi \to B\phi$
Eqs. (2.2.2), (2.2.3)	*Knowledge entails* ***true*** *belief*: $\mathscr{K} \subseteq \beta$	Axiom (**FL1**), Eqs. (4.7.1) (4.7.2)	$\mathbb{M} \models K\phi \to (\phi \land B\phi)$
Prop. 2.2.3, Eq. (2.2.4)	*Knowledge entails* ***possibly grounded*** *belief*: $\mathscr{K} \subseteq$ ᚹ	Axiom (**FL1**)	$\mathbb{M} \models K\phi \to \diamondplus B\phi$
Prop. 2.2.4, Eq. (2.2.5)	***Fully grounded*** *belief entails knowledge*: $\flat \subseteq \mathscr{K}$	Axiom (**FL2**)	$\mathbb{M} \models (\diamondplus B\phi \land \boxminus(B\phi \to \diamondplus B\phi)) \to K\phi$

	Philosophy		Logic
	Belief Basis		
Def. 2.2.3	*The deductive closure of a belief basis is the belief base*: $\overline{\mathfrak{B}} = \mathscr{B}$	Thrm. 4.6.8	*Deductive closure theorem*: For any world $\langle b, \mathfrak{B}\rangle$ in any model $\mathbb{M}$, $\overline{\mathfrak{B}} \subseteq \mathscr{B}_{\langle b,\mathfrak{B}\rangle}$; if $\mathfrak{B}$ is finite, then $\overline{\mathfrak{B}} = \mathscr{B}_{\langle b,\mathfrak{B}\rangle}$
Def. 2.2.4(a)	$\mathfrak{C}$ is said to be *grounded* if and only if for every $\psi \in \mathscr{L}_{\mathsf{L}}$, if $\psi \in \mathfrak{C}$, then ψ is True in the relevant circumstances	§4.4.3	R_+ relation accesses grounded subsets of a belief basis at $\langle b, \mathfrak{B}\rangle$
Def. 2.2.4(b)	An agent is said to have *belief in ϕ which is possibly grounded* if and only if there is some $\mathfrak{C} \subseteq \mathfrak{B}$ which is grounded and $\phi \in \overline{\mathfrak{C}}$	§4.4.3	$(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondplus B\phi$
Prop. 2.2.2	*The justification principle*: if an agent has a possibly grounded belief in ϕ, then ϕ is True in the current circumstances	Thrm. 4.6.2	*The justification theorem*: $\models \diamondplus B\phi \to \phi$
Def. 2.2.6(a)	$\mathfrak{C}$ is said to be *specious* if and only if there is some $\psi \in \mathfrak{C}$ such that ψ is False in the relevant circumstances	§4.4.4	R_- relation accesses specious subsets of a belief basis at $\langle b, \mathfrak{B}\rangle$
Def. 2.2.6(b)	An agent is said to have *belief in ϕ which is possibly specious* if and only if there is some $\mathfrak{C} \subseteq \mathfrak{B}$ which is specious and $\phi \in \overline{\mathfrak{C}}$	§4.4.4	$(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondminus B\phi$
Def. 2.2.7	An agent with a belief basis $\mathfrak{B}$ is said to have a *fully grounded* belief in ϕ if and only if (a) ϕ is possibly grounded (b) for every specious $\mathfrak{C} \subseteq \mathfrak{B}$ such that $\phi \in \overline{\mathfrak{C}}$, then there is a grounded subset $\mathfrak{D} \subseteq \mathfrak{C}$ such that $\phi \in \overline{\mathfrak{D}}$	Axiom (**FL2**)	$(\mathbb{M}, \langle b, \mathfrak{B}\rangle) \models \diamondminus B\phi \wedge \boxminus(B\phi \to \diamondplus B\phi)$

I therefore have a complete lexicon of my entire philosophy on the knowledge of ideal agents in terms of my semantics. I personally do not see much distinction between reasoning philosophically regarding this topic and reasoning with my semantics.

5.2 Further Work

In this section I present some further work I foresee for my research.

5.2.1 Soundness and Completeness

In this section I present what I think might be a complete axiomatization of a restriction of my semantics. If I completely disregard R_K, and deal purely with $B, \boxplus$, and $\boxminus$, I conjecture the following characterizes my semantics for FL:

Modus Ponens Rule

$$\frac{\vdash \phi,\ \vdash \phi \to \psi}{\vdash \psi}$$

(FL′1) All propositional tautologies

(FL′2) $B(\phi \to \psi) \to (B\phi \to B\psi)$

(FL′3) $\boxplus(\phi \to \psi) \to (\boxplus\phi \to \boxplus\psi)$

(FL′4) $\boxminus(\phi \to \psi) \to (\boxminus\phi \to \boxminus\psi)$

(FL′5) $B\phi \leftrightarrow (\Diamond\!\!\!+ B\phi \vee \Diamond\!\!\!- B\phi)$

(FL′6) $\neg \boxplus \neg\phi \leftrightarrow \Diamond\!\!\!+\phi$

(FL′7) $\neg \boxminus \neg\phi \leftrightarrow \Diamond\!\!\!-\phi$

(FL′8) $\Diamond\!\!\!+ B\phi \to \phi$

(FL′9) $\neg \boxminus \bot \to (\boxminus\phi \to \phi)$

(FL′10) $\boxminus\bot \to (\boxplus\phi \to \phi)$

(FL′11) $\neg \boxminus \bot \vee \neg \boxplus \bot$

(FL′12) $\boxminus\phi \to \boxminus \boxminus \phi$

(FL′13) $\boxplus\phi \to \boxminus \boxplus \phi$

(FL′14) $\boxplus\phi \to \boxplus \boxplus \phi$

(FL′15) $\boxplus \boxminus \bot$

(FL′16) $\boxminus \neg \boxminus \bot$

Necessitation Rules

$$\frac{\vdash \phi}{\vdash B\phi}$$

$$\frac{\vdash \phi}{\vdash \boxplus\phi}$$

$$\frac{\vdash \phi}{\vdash \boxminus\phi}$$

For the class of models that are *closed under union of grounds*:

(FL′17) $\diamondplus B(\phi \to \psi) \to (\diamondplus B\phi \to \diamondplus B\psi)$

For the class of models that are *Cartesian*:

(FL′18) $\neg \boxminus \bot$

(FL′19) $B \boxplus B\phi \leftrightarrow \boxplus B\phi$

(FL′20) $\boxplus \boxplus B\phi \to \boxplus B\phi$

(FL′21) $\boxminus \boxplus B\phi \to \boxplus B\phi$

(FL′22) $\diamondplus \boxplus B\phi \leftrightarrow \boxplus B\phi$

(FL′23) $B \neg \boxplus B\phi \leftrightarrow \neg \boxplus B\phi$

(FL′24) $\boxplus \neg \boxplus B\phi \leftrightarrow \neg \boxplus B\phi$

(FL′25) $\boxminus \neg \boxplus B\phi \leftrightarrow \neg \boxplus B\phi$

(FL′26) $\diamondplus \neg \boxplus B\phi \leftrightarrow \neg \boxplus B\phi$

Cartesian Necessitation Rule

$$\frac{\vdash \phi}{\vdash \diamondplus\phi}$$

The proof of the completeness of this axiomitization is utterly beyond me; I cannot use Kripke's canonical models due to the uniqueness of my semantics. A reader unfamiliar with proof theory in modal logic should consult a text on the subject, such as Blackburn et al. (2001). I further conjecture that FL' is *decidable*, however I am not sure of the complexity. I optimistically hope it is NP, but this is questionable. A reader unfamiliar with computational complexity theory of decision processes for logics may also be interested in Blackburn et al. (2001).

Furthermore, I feel this research could some day form a bridge between the analysis of epistemic logic put forward by computer scientists and the analysis put forward by philosophers. This is because on the one hand I feel that $\boxplus B$, when $\mathbb{M}$ is Cartesian, behaves a lot like the K that computer scientists commonly study; this can be seen through the following lemma, which I state without proof:

LEMMA 5.2.1 If Γ is a set of propositions that contains all of the ordinary axioms and the Cartesian axioms, and all of the rules of inferences (including Cartesian necessitation) hold, then:

(*i*) $\Gamma \vdash \boxplus B\phi \to \phi$

(*ii*) $\Gamma \vdash \boxplus B\phi \to B\phi$

(*iii*) $\Gamma \vdash \boxplus B\phi \to \boxplus B \boxplus B\phi$

(*iv*) $\Gamma \vdash \neg \boxplus B\phi \to \boxplus B \neg \boxplus B\phi$

Thus, it can be seen that $\boxplus B$ acts like **S5**, just like the knowledge commonly studied in computer science. On the other hand, if Γ contains all the axioms then $\oplus B$ also behaves like a modal operator which obeys axiom **KT**. This is more like the knowledge operator commonly studied in epistemic logic, however a reader will note that just because $\oplus B\phi$ is true in some instance does not mean the *sufficient* condition I outlined. Indeed, no necessary and sufficient conditions for philosophical knowledge exist, so $\oplus B$ merely serves as at best a crude approximation.

5.2.2 Experience About Beliefs & Multiple Agents

In this section I talk about deficits that my semantics suffers that I feel can be over come. Recall that beliefs about beliefs cannot be modeled in belief basies; furthermore, it seems like the same sorts of paradoxes emerge when one tries to model multiple agents reasoning about one another. I feel that the same trick can be used to remedy both situations.

The trick is to give multiple grammars, as I did in ordinary semantics for FL. Consider the following three grammars:

DEFINITION 5.2.1
Let $\mathscr{L}^0_{\mathsf{FL}}(\mathbf{P})$ be the smallest set such that:

(*a*) $\mathbf{P} \subseteq \mathscr{L}^0_{\mathsf{FL}}(\mathbf{P})$

(*b*) If $\{\phi, \psi\} \subseteq \mathscr{L}^0_{\mathsf{FL}}(\mathbf{P})$, then $(\phi \wedge \psi) \in \mathscr{L}^0_{\mathsf{FL}}(\mathbf{P})$

(*c*) If $\phi \in \mathscr{L}^0_{\mathsf{FL}}(\mathbf{P})$, then $\neg\phi \in \mathscr{L}^0_{\mathsf{FL}}(\mathbf{P})$

Let $\mathscr{L}^1_{\mathsf{FL}}(\mathbf{P})$ be the smallest set such that:

(*a*) $\mathscr{L}^0_{\mathsf{FL}}(\mathbf{P}) \subseteq \mathscr{L}^1_{\mathsf{FL}}(\mathbf{P})$

(*b*) If $\{\phi, \psi\} \subseteq \mathscr{L}^1_{\mathsf{FL}}(\mathbf{P})$, then $(\phi \wedge \psi) \in \mathscr{L}^1_{\mathsf{FL}}(\mathbf{P})$

(*c*) If $\phi \in \mathscr{L}^1_{\mathsf{FL}}(\mathbf{P})$, then $\{\neg\phi, K^1_1\phi, K^1_2\phi, B^1_1\phi, B^1_2\phi, \boxplus^1_1\phi, \boxplus^1_2\phi, \boxminus^1_1\phi, \boxminus^1_2\phi\} \subseteq \mathscr{L}^1_{\mathsf{FL}}(\mathbf{P})$

Let $\mathscr{L}^2_{\mathsf{FL}}(\mathbf{P})$ be the smallest set such that:

(a) $\mathscr{L}^1_{\mathsf{FL}}(\mathbf{P}) \subseteq \mathscr{L}^2_{\mathsf{FL}}(\mathbf{P})$

(b) If $\{\phi, \psi\} \subseteq \mathscr{L}^2_{\mathsf{FL}}(\mathbf{P})$, then $(\phi \wedge \psi) \in \mathscr{L}^2_{\mathsf{FL}}(\mathbf{P})$

(c) If $\phi \in \mathscr{L}^2_{\mathsf{FL}}(\mathbf{P})$, then $\{\neg\phi, K^1_1\phi, K^1_2\phi, B^1_1\phi, B^1_2\phi, \boxplus^1_1\phi,\ \boxplus^1_2\phi,\ \boxminus^1_1\phi,\ \boxminus^1_2\phi,\ K^2_1\phi,\ K^2_2\phi,\ B^2_1\phi, B^2_2\phi, \boxplus^2_1\phi, \boxplus^2_2\phi, \boxminus^2_1\phi, \boxminus^2_2\phi\} \subseteq \mathscr{L}^2_{\mathsf{FL}}(\mathbf{P})$

... the basic idea being that $K^i_j\phi$ corresponds to "level i" knowledge for agent j, and so on. Worlds now look like $\langle b, \vec{\mathfrak{B}}_1, \vec{\mathfrak{B}}_2\rangle$, where $\vec{\mathfrak{B}}_i = \langle \mathfrak{B}^1_i, \mathfrak{B}^2_i\rangle$, and $\mathfrak{B}^1_i \subseteq \mathscr{L}^0_{\mathsf{FL}}$ and $\mathfrak{B}^2_i \subseteq \mathscr{L}^1_{\mathsf{FL}}$. Now, relations are defined in a more complicated manner, with $R^1_{B,1}$, $R^1_{B,2}$,$R^2_{B,1}$, $R^2_{B,2}$ etc, where for instance:

Proposition 5.2.1 $R^i_{B,j}(\langle b, \vec{\mathfrak{B}}_1, \vec{\mathfrak{B}}_2\rangle, \langle c, \vec{\mathfrak{C}}_1, \vec{\mathfrak{C}}_2\rangle)$ if and only if $(\mathbb{M}, \langle c, \vec{\mathfrak{C}}_1, \vec{\mathfrak{C}}_2\rangle) \models \mathfrak{B}^i_j$

I hold that this forms a paradox free semantics, which allows for me to model systems where $\mathbb{M} \models B^2_j\phi \to B^1_j\phi$, and $\mathbb{M} \models B^1_j\phi \to B^2_jB^1_j\phi$, and therefore $\mathbb{M} \models B^2_j\phi \to B^2_jB^2_j\phi$. Other communication between beliefs may also be possible; this system promises a potentially expressive means for reasoning about agents reasoning about each other.

5.3 Conclusion

I hope this paper provides a novel approach to epistemic logic; the concepts I have developed in this paper may be applicable to other branches of epistemic logic, as the business of modeling agents thinking about their world is easily suited to having the concept of a *basis of belief* applied; if one permits such structures to induce accessibility relations, then one may arrive at a modal logic. My greatest wish is that the ideas I have presented here find use to someone with like minded interests.

References

S. Artemov and E. Nogina. Introducing justification into epistemic logic. *J. Log. and Comput.*, 15(6):1059–1073, 2005a. ISSN 0955-792X. doi: 10.1093/logcom/exi053. URL http://dx.doi.org/10.1093/logcom/exi053.

S. Artemov and E. Nogina. On epistemic logic with justification. In *TARK '05: Proceedings of the 10th conference on Theoretical aspects of rationality and knowledge*, pages 279–294, Singapore, Singapore, 2005b. National University of Singapore. ISBN 981-05-3412-4.

P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*. Number 53 in Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 2001. doi: 10.2277/0521527147.

G. Boolos. *The Logic of Provability*. Cambridge University Press, 1993. ISBN 9780521483254.

M. Clark. Knowledge and grounds: A comment on mr. gettier's paper. *Analysis*, 24(2):46–48, dec 1963. ISSN 0003-2638. URL http://links.jstor.org/sici?sici=0003-2638%28196312%2924%3A2%3C46%3AKAGACO%3E2.0.CO%3B2-H.

H. Enderton. *A Mathematical Introduction To Logic*. Academic Press New York, 1972. ISBN 0122384520.

R. Fagin, J. Y. Halpern, M. Y. Vardi, and Y. Moses. *Reasoning about knowledge*. MIT Press, Cambridge, MA, USA, 1995. ISBN 0-262-06162-7.

E. L. Gettier. Is justified true belief knowledge? *Analysis*, 23(6):121–123, jun 1963. ISSN 0003-2638. URL http://links.jstor.org/sici?sici=0003-2638%28196306%2923%3A6%3C121%3AIJTBK%3E2.0.CO%3B2-1.

K. Gödel. *On Formally Undecidable Propositions of Principia Mathematica and Related Systems*. Dover Publications, April 1992. ISBN 0486669807.

J. Y. Halpern. Should knowledge entail belief? *Journal of Philosophical Logic*, 25(5):483–494, 1996. URL http://dx.doi.org/10.1007/BF00257382.

J. Y. Halpern. A Theory of Knowledge and Ignorance for Many Agents. *J Logic Computation*, 7(1):79–108, 1997. doi: 10.1093/logcom/7.1.79. URL http://logcom.oxfordjournals.org/cgi/content/abstract/7/1/79.

J. Y. Halpern. Set-theoretic completeness for epistemic and conditional logic. *Annals of Mathematics and Artificial Intelligence*, 26(1-4):1–27, 1999. ISSN 1012-2443.

J. Y. Halpern and Y. O. Moses. Towards a theory of knowledge and ignorance: A preliminary report. In *Proceedings of the AAAI Workshop on Non-Monotonic Reasoning*, pages 125–143, New Paltz, NY, 1984. Morgan Kaufmann. URL citeseer.ist.psu.edu/halpern85towards.html.

J. Hintikka. *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Cornell University Press, 1962. ISBN 0801401879.

B. P. Kooi. Probabilistic dynamic epistemic logic. *J. of Logic, Lang. and Inf.*, 12(4):381–408, 2003. ISSN 0925-8531. doi: 10.1023/A:1025050800836. URL http://dx.doi.org/10.1023/A:1025050800836.

S. Kraus and D. Lehmann. Knowledge, belief and time. *Automata, Languages and Programming*, pages 186–195, 1986. URL http://dx.doi.org/10.1007/3-540-16761-7_68.

H. Kyburg and C. Teng. *Uncertain Inference*. Cambridge University Press, 2001. ISBN 0521001013.

W. Lenzen. *Recent Work in Epistemic Logic*. North-Holland, 1978. ISBN 9519505407.

J.-J. C. Meyer and W. van der Hoek. *Epistemic Logic for AI and Computer Science*. Cambridge University Press, New York, NY, USA, 1995. ISBN 052146014X.

G. E. Moore and W. H. Shaw. *Ethics*, volume 1. Oxford Scholarship Online Monographs, 25 August 2005. URL http://www.ingentaconnect.com/content/oso/1116139/2005/00000001/00000001/art00000.

Plato. *Meno*, volume 1643 of *Project Gutenberg*. Project Gutenberg, P.O. Box 2782, Champaign, IL 61825-2782, USA, 1999a. URL ftp://uiarchive.cso.uiuc.edu/pub/etext/gutenberg/etext99/1meno10.zip.

Plato. *Parmenides*, volume 1687 of *Project Gutenberg*. Project Gutenberg, P.O. Box 2782, Champaign, IL 61825-2782, USA, 1999b. URL ftp://uiarchive.cso.uiuc.edu/pub/etext/gutenberg/etext99/prmds10.zip.

René Descartes. *Meditations on First Philosophy*. Wright State University, 1996. URL http://www.wright.edu/cola/descartes/intro.html.

A. Schopenhauer. Ueber di vierfache wurzel des satzes vom zureichenden grunde. In *Two Essays*, chapter vii, sec. 41, page p. 166. G. Bell, 1889.

W. van der Hoek. Systems for knowledge and beliefs. *Logics in AI*, pages 267–281, 1991. URL http://dx.doi.org/10.1007/BFb0018447.

W. van der Hoek and E. Thijsse. Honesty in partial logic. In J. Doyle, E. Sandewall, and P. Torasso, editors, *KR'94: Principles of Knowledge Representation and Reasoning*, pages 583–594. Morgan Kaufmann, San Francisco, California, 1994. URL citeseer.ist.psu.edu/article/vanderhoek93honesty.html.

H. van Ditmarsch, W. van der Hoek, and B. Kooi. Concurrent dynamic epistemic logic, 2003. URL http://citeseer.ist.psu.edu/vanditmarsch03concurrent.html.

F. Voorbraak. Generalized kripke models for epistemic logic. In *TARK '92: Proceedings of the 4th conference on Theoretical aspects of reasoning about knowledge*, pages 214–228, San Francisco, CA, USA, 1992. Morgan Kaufmann Publishers Inc. ISBN 1-55860-243-9.

L. Zagzebski. The inescapability of gettier problems. *The Philosophical Quarterly*, 44(174):65–73, jan 1994. ISSN 0031-8094. URL http://links.jstor.org/sici?sici=0031-8094%28199401%2944%3A174%3C65%3ATIOGP%3E2.0.CO%3B2-R.