SVEUČILIŠTE U RIJECI

FILOZOFSKI FAKULTET U RIJECI

ODSJEK ZA FILOZOFIJU

ANTE DEBELJUH

U OBRANU NORMALNE EPISTEMIČKE LOGIKE

DOKTORSKI RAD

Rijeka, 2025.

SVEUČILIŠTE U RIJECI

FILOZOFSKI FAKULTET U RIJECI

ODSJEK ZA FILOZOFIJU

ANTE DEBELJUH

U OBRANU NORMALNE EPISTEMIČKE LOGIKE

DOKTORSKI RAD

Mentor: dr.sc. Nenad Smokrović, prof. emeritus

Rijeka, 2025.

UNIVERSITY OF RIJEKA

FACULTY OF HUMANITIES AND SOCIAL SCIENCES

DEPARTMENT OF PHILOSOPHY

ANTE DEBELJUH

IN DEFENCE OF NORMAL EPISTEMIC LOGIC

DOCTORAL THESIS

Rijeka, 2025.

Mentor doktorskog rada: titula, ime i prezime

Doktorski rad obranjen je dana ____

u/na ____ ,

pred povjerenstvom u sastavu:

1. ____

2. ____

3. ____

For Ema,

mum and dad,

and, of course, Nenad

SUMMARY

The problem that this dissertation addresses is strictly defining the relationship between an epistemological and doxastic theory and a formal theory that supports it. When we state that a formal theory supports an epistemology, in essence we are stating that the axioms and the inferential structures that the formalism offers validate the theoretical claims within the informal theory that it attempts to model.

The final product of such an endeavour is a formal model, comprising a formal theory of knowledge and belief. The model that I am proposing is established in order to model a contemporary account of verificationist epistemology, which will be displayed within a formal setting of Distributed Systems Models. It comprises three distinct layers defined as formal structures; (1) the base structure that offers a formal verificationist theory of meaningfulness, (2) the epistemic and doxastic structure that defines the notions of knowledge and belief through two distinct logical systems, S4.2 for knowledge and CDL for belief, and finally (3) the superstructure that dynamises knowledge by implementing the notion of algorithmic, i.e. computable knowledge. I will show that the model is validated by a Kripke structure, an upshot of which is maintaining normality of the system. The normality of the modal logics that are used guarantees that the system remains sound and complete.

Furthermore, I will discuss possible solutions to (or trivialisations of) some problems that arise within the discussion, such as the problem of logical omniscience for normal systems and Fitch's paradox with original arguments which further constitute my contribution to the discussion. Finally, I will address the introduced systems in terms of their metalogical properties.

Table of Contents

CHAPTER I – INTRODUCTION

The concepts of knowledge and belief have been contentious for the better part of the last two and a half millennia, and have been discussed in the philosophical discourse, mainly in their central domain of epistemology, but also in adjacent fields of philosophy of logic, ethics, philosophy of art, philosophy of mind, etc. Although the two concepts have been central in the history of philosophy since Plato's seminal work Theaetetus, in which he discussed them through his dialogues, they were present in the discourse even earlier with authors such as Parmenides (On Nature), Xenophanes (preserved in the works of Sextus Empiricus, Adversus Mathematicos VII.49), and Heraclitus (preserved in the works of Sextus Empiricus, Adversus Mathematicos VII.133). There was an evident need for a clear definition, as the concepts were seemingly used meaningfully in everyday discourse, but there was no universal convergence on what they *exactly* pertain to. Plato's three-part definition of knowledge that stated knowledge to comprise a *justified, true belief* was considered mainstream until the second half of the twentieth century when Edmund Gettier wrote the paper entitled "Is Justified True Belief Knowledge?" (1963). He was certainly not the first to come up with counterexamples to the definition, as Plato himself was already aware of some of its shortcomings, but Gettier's paper popularised the problem to such an extent that an entire branch of philosophy (often called the post-Gettier epistemology) was developed in the endeavour to finally arrive at a definition of knowledge that would be capable of bypassing the pitfalls of the Plato's original definition. In the Plato's version belief was taken to be primitive, and this idea was later adapted to the contemporary language to be represented as an *intentional mental state* that pertains to some state of affairs.

A year earlier than Gettier's paper on counterexamples of the Plato's definition there was another development in the field of epistemology virtually regarding the same issue, but from a vastly different angle. Jaakko Hintikka (1962), a Finnish logician and philosopher has developed a mathematical description of the two notions using the instruments of formal modal logic that attempted to display the definitions and interrelations of knowledge and belief in an abstract setting. This attempt would render knowledge and belief formally explicable in order to get a complete and unambiguous grasp on what exactly they are, along with all the consequences that followed from their conceptions. This did not mean in any reasonable sense that there was only one formal representation of them (which would most probably result on a final convergence on the issue of their interdefinability), but only that we had the tools to define clearly, while having a proof theory that would account for every consequence of adopting some definition of them. The relationship between a formal model and an informal theory was discussed at lengths by authors such as Williamson (2000), Lewis (1996), and even Hintikka himself (1962).

Each coherent epistemological theory could be accounted for with this newly founded apparatus by means of combining and manipulating axiomatic schemas and rules of inference for the modal system that captures the theory in order to get its formal explication. The formal explication of the theory is then said to *model* a philosophical position, and can, for instance, be applied on capturing descriptions of various epistemic and doxastic configurations in the fields of computer science and the development of artificial intelligence. This is only to say that a model supported by systems of epistemic and doxastic logics can be used as normative instruments for they can instruct an agent how to infer and behave in certain abstract situations that are sensitive to knowledge, belief, and action of the agents in question. The kind of analysis that discusses applications of epistemic and doxastic logics as normative standards for systems in computer sciences and the development of AI can be found in Fagin et al. (1995), Bordini et al. (2006), and Wooldridge (2009).

THE THESIS AND ITS CONTRIBUTION TO THE DISCUSSION

The problem that this dissertation addresses is strictly defining the relationship between an epistemological and doxastic theory and a formal theory which should support it. The final product of such an endeavour is a formal model. The central theses of this dissertation can, hence, be stated as following;

(1) *A modal account* of the verificationist position is, in my view, the ideal framework for modelling semantic meaningfulness, knowledge, belief, and action of resource-bound agents in a dynamic multi-modal and multi-agent abstract setting.

(2) It is possible, and even more so – optimal, to model such a verificationist account with normal epistemic and doxastic logics.

There is no way to meaningfully claim that a model (of any theory) is *correct* in terms of capturing some phenomena, so the central point of the dissertation deals with the issue of *framing and application* of normal systems to the relevant notions in order to make them operative for establishing explications, explanations, and anticipations of the behaviour of the agents that the system pertains to.

Furthermore, I will discuss possible solutions to some problems, such as the problem of logical omniscience, Fitch's paradox of knowability, and the problem of algorithmic inference with original approaches to their potential solutions, which further constitute the contribution of this dissertation to the present discussion. I have chosen to break this dissertation up into nine discrete chapters, most of which might be capable of existing as a standalone texts pertaining to various philosophical and logical areas needed for the completion of the model.

The model that I am proposing is established in order to model a contemporary account of verificationist epistemology, which will be displayed within a formal setting of Distributed Systems Models. It comprises three distinct layers defined as formal structures; (1) the base structure that offers a formal verificationist theory of meaningfulness, (2) the epistemic and doxastic structure that defines the notions of knowledge and belief through two distinct logical systems, S4.2 for knowledge and CDL for belief, and finally (3) the superstructure that dynamises knowledge by implementing the notion of algorithmic, i.e. computable knowledge. I will show that the model is validated by a Kripke structure, an upshot of which is maintaining normality of the system. The normality of the epistemic and doxastic logics that are used guarantees that the system remains sound, complete, compact, and canonical.

It was never a plan to devise original logical systems whose axioms support the specific verificationist epistemology I have constructed for this dissertation, I believe that such hybrid epistemic-doxastic structure, which uses S4.2 logic for knowledge (also advocated by Stalnaker and Lenzen) and CDL for belief (also advocated by Negri and Pavlović) has not yet been considered for modelling verificationist epistemic positions. I further attempt to show that the logic knowledge S4.2 is not only compatible with a realist epistemology, as was the case in Stalnaker's and Lenzen's respective models, but with an antirealist one as well if we were to adapt the philosophical interpretation of the axiom of factivity to pertain to demonstrable or verifiable truths, as opposed to truths simpliciter (as is the case in the realist interpretation of the S4.2 system). As far as I am aware, no attempts were so far made to model a verificationist epistemology with an S4.2 epistemic logic, so this can potentially be viewed as a further original contribution of the dissertation to the discussion.

As for the epistemic theory that I have opted for in terms of this dissertation, I propose a revised account of the verificationist epistemology which defines the criterion of meaningfulness in modal terms. Its dynamic structure allows for the notion of meaningfulness to be understood as an object of discovery – we discover that a statement (a syntactic object) is a bearer of a proposition (a logical object) by being capable of cognitively constructing a situation in which it obtains, and furthermore, discriminating such a situation from one in which it does not. This requires us to be apt to construct an *in principle* test that measures the variable in adequate surroundings that makes it true in one situation and false in another. This idea relies on the view that everything apart from the mathematical and logical truths that we are capable of conceptualising and consequently understanding ought to be empirically constructible or accessible, even if only in principle.

A good comparison between a theory that examplifies this and one that does not was proposed to me by prof. Berčić after reading the first draft of the dissertation, and pertains to Rudolf Carnap's and Thomas Nagel's disagreement on the matter. The view akin to the one I am offering here might be found when analysing Rudolf Carnap's article The

Elimination of Metaphysics Through Logical Analysis of Language (1932). Carnap states:

"If someone were to assert the existence of a being which is neither spatial nor temporal, which cannot be experienced by any of the senses, and which nevertheless is supposed to influence the course of events in the world — we could not say that his assertion is false, but rather that it is nonsensical. Even if someone — say the devil or God himself — were to tell us that such a being exists, we would still not know what it is that has been asserted." (Carnap, 1932)

Opposing Carnap on this issue, Nagel in his monograph *The View from Nowhere* (1986) states in Chapter 1:

"We cannot claim to occupy a position outside of ourselves, to see the world from no point of view at all. But neither can we deny that there is a reality that transcends our own perspective and that much of it must remain beyond our understanding." (Nagel, 1986)

He continues on this matter in Chapter 4 by stating:

"There may be aspects of reality that human beings are simply not capable of understanding, because of the limitations of our concepts and our form of perception." (Nagel, 1986)

With thanks to prof. Berčić for the comment, I push against Nagel's notion of *unknown unknowns*. It should be further noted that Nagel's notion of *unknown unknowns* goes even beyond that what the phrase originally suggests – his view on this matter should be understood as accepting the existence of *unknown unknowables*, or rather the idea that there exist truths in the world that we cannot even in principle capture by our cognitive apparatus. Here I opt for a Carnapian conception of what is cognitively operative within an epistemic theory – statements that are to be the candidates for playing a role in our cognitive economy (or rather be considered meaningful at all) are the ones that are at least in principle testable or constructible. In other words, if we have no cognitive means to construct a possible test, even in optimal epistemic circumstances, to discriminate between a situation in which the statement is true from the one in which it is not, then the statement should be dismissed as meaningless.

This might appear to be not disimilar to Dummett's (1978) semantic interpretation of the verificationist epistemology, however, my view on this matter is that such epistemologies are not necessarily to be supported by a logical enterprise weaker than classical as was the case with Dummett's intuitionistic approach to the matter. As I have stated, and as the title of the work suggests, I intend to show that there exists a philosophical reading of normal epistemic and doxastic logics S4.2 and CDL, respectively, that supports this Carnapian approach.

Even though I have attempted to make this dissertation as beginner-friendly as I could, so some of its parts could serve as introductory texts to the domain of formal epistemology, some background knowledge of propositional modal calculus is welcome. Furthermore, I have opted not to venture into too much detail when introducing some systems that are not of central relevance to the model I am constructing, but have included various additional literature pertaining to such systems for the curious readers. I have also attempted to homogenise both formal and informal vocabulary used throughout the dissertation. This is because the relevant authors in this field come from a plethora of academic backgrounds such as computer science, mathematical logic, philosophy, etc., and they often opt for terminology from their 'home' field, which results in a widely non-unified vocabulary on every front. The other point of importance that I want to address here is the framing of the problems across various fields that deal with formal epistemology. We will observe, especially in the most interdisciplinary topics such as the Distribution Systems Models, that there exists a theoretical divergence in the formal model framing from different authors. All of their statements and definitions are perfectly translatable from one frame into another without anything of importance being lost in the process, but a new reader might encounter more than a few headscratchers when first attempting to make sense of the field. As far as the formal vocabulary goes (in refence to logical terms of the relevant systems), the divergence remains quite unhomogenised across the field. When paraphrasing the papers and position of various authors, I have attempted to consolidate the terminology to a single standard, as I believe that otherwise the dissertation would become unreadable and impossible to make sense of. That being said, when dealing with such paraphrases, I have tried my best to change as little as necessary in order to respect the source material. I hope that the readers find my chosen nomenclature clear and intuitive in order for the point to get across. Finally, before getting to the introduction to the topic and the structure of the dissertation, I want to say that the field of formal epistemology is absolutely vast and that no answers given can be considered conclusive in any reasonable sense of the word, as there are (countably) infinitely many perspectives on any discussed problem.

When we state that a formal theory supports an epistemology, in essence we are stating that the inferential structures that the formalism offers validates the theoretical claims within the informal theory that it attempts to model. To start with, several important issues ought to be addressed in order to make this discussion clear. Firstly, it is important to understand that every scientific and philosophical theory has a logic that is implied in its structure. This means that the structures that support the theoretical concepts that the system uses are defined by the underlying logic. It ought to be said that not every statement of the theory ought to be formally represented within the logic that it uses, it is mainly of importance that there does not exist any tension between the informal theory and the formalism.

In the first substantive chapter of the dissertation I offer an introduction to the concept of modalities and intensional semantics. I offer a brief overview of the historic motivations for developing modal systems and tackle several example of some uses of modal concepts in natural language. I mainly focus on the examples that possess relevance for the formal systems I will develop in further chapters of the book. Within the chapter I have also chosen to tangentially discuss the types of conditionals that are recognised as grammatical categories in natural language, along with their treatments in terms of intensional analysis

(Lewis, 1963; Kratzer, 1991; Veltman, 1996). As their role is central to the understanding of modality and formalisms that support them, I saw fit to introduce them in this chapter.

The second chapter of the dissertation deals with the introduction to epistemic and doxastic logics, in which I present the systems' inferential inventory, axioms, and language. I discuss three epistemic and doxastic systems; S4.2, S5, and KD45 as a sort of a case study in order for the reader to get a better grasp of how the systems are internally arranged and defined. Furthermore, I talk about the infrastructure of the systems from the perspective of metalogic and offer definitions of frames and models that validate the proposed axiomatic schemas (Stalnaker, 2019). Within this chapter I will briefly tackle the issue of calibrating formalisms to epistemological and doxastic theories and will talk about the issue of 'translatability' of informal language into a formal one. Finally, I talk about the relationship between the axioms of the system and the restrictions on the accessibility relations, which define the systems' inferential apparatus (Chellas, 1980; Garson, 2006).

As the systems that were tackled so far are all static, in the next chapter of the dissertation, I offer an overview of a dynamic epistemic models that were mainly developed for the purposes of modelling structures within the domain of computer science – the Distributed Systems Models (Stalnaker, 1999). The dynamisation of epistemic and doxastic logics opens the door for solutions to some quite contentious problems in the field, such as the problem of logical omniscience. Since this problem is usually tightly bound to the analysis of normal logics and their dynamisation, here I propose a clarificatory portion of the text that should make sure that the problem is adequately fleshed out, so we can observe the possible solutions that are on the table. In continuation of the chapter I introduce the notion of algorithmic knowledge, which is suited for the discourse of epistemic externalism on which Distributed Systems Models rely (Halpern et al. 1994). At the end of the chapter I show how the dynamisation of the system that was generated by the introduction of algorithmic knowledge gives way to expanding the intended domain of application of such models from highly idealised agents to our ordinary knowers. Within the final part of the chapter, I show how we are able to express algorithmic knowledge within the modal plane by using a separate intensional operator, which will serve as an instrument of differentiating between what is in principle inferable within the system (knowledge simpliciter) and what is de facto calculable by the agents of the system (algorithmic knowledge). I will also show why this works as a possible solution to the logical omniscience problem regarding the modality of knowledge (Stalnaker, 1999).

The next chapter works in a similar manner, but the intended modality that I display in the dynamic setting is belief. I will approach the dynamisation of belief from a different perspective. First I will elaborate on the dynamic turn of epistemic and doxastic logics through the DEL and PAL systems (van Benthem, 2007), then I will show how we can approach the dynamics of belief from a different angle – the system of conditional belief entitled Conditional Doxastic Logic (CDL) (Negri and Pavlović, 2023). This system is not dynamic in itself, however, it is capable of capturing the dynamics of belief revision through a set of relations between the body of knowledge and its consequence in form of a belief. It defines the structure of what propositions we are justified to accept once our set of known propositions is changed in some respect. This system is axiomatically very similar to the static logic of belief (KD45) that we have observed in the chapter on epistemic and doxastic logics, however, it models the basic postulates of the AGM theory of belief revision (Alchurron et al., 1985) through the mentioned set of conditional relations. AGM theory itself lacks in the department of explicating rules of inference when describing the shift in the belief states, but the framework that it provides through the set of postulates of belief revision appears to be quite useful when developing this kind of analysis. Furthermore, I will show that the since the system that we have observed for dynamising the models of belief (CDL) is axiomatically equivalent to the KD45 system with added and appropriated modality for conditional belief, our basic S4.2 system that was used for modelling knowledge will be equally adapted to CDL as it was the case with KD45. All appropriate properties of the system remain unchanged. This being said, the systems of conditional belief are explicated more elegantly and naturally in the context of neighbourhood frames as opposed to Kripke frames. I will provide an introduction to neighbourhood frames, with emphasis of defining accessibility relations through the language of set theory, i.e. through '*truth sets*' and membership relations. Even with this being the case, there exists proof (Pacuit, 2017) that that each neighbourhood frame that is augmented by two conditions, viz. validating the distribution axiom (K axiom) and the rule of necessitation (Nec), corresponds to a linear Kripke frame with binary accessibility relations. Validating these two conditions guarantees that the system remains normal.

The penultimate substantive chapter of the dissertation finally arrives at the philosophical perspective on this project, where I opt for an epistemological theory that the model we have been observing validates. I opted for a semantic verificationist account that takes what I consider essential from the verificationist enterprise and I show how to calibrate the formal theory for the purpose of establishing a complete model. In this chapter I will discuss some assumptions that ought to be made for a verificationist epistemology not to collapse under its own weight and offer what I hope is a fresh perspective on devising such a theory. Verificationism as an epistemological theory is specific in the sense of developing a very intricate and close relationship between semantics and epistemology. The flavour of verificationism that I opted to defend in this chapter relies on a simple idea of grounding meaningfulness on the intensionally motivated idea of constructing cognitive conditions through the set of accessible worlds. Some similar ideas (but not the central one) can be found in the monograph *Minimal Verificationism* (Haas, 2015). I reject most of his conclusions as we attempt to defend two rather different verificationist accounts, but we converge in the method of using modal apparatus for defining semantic and cognitive meaningfulness. In the chapter I also address a couple of possible extensions of the model such as introducing common and distributed knowledge that account for a verificationist framing of the model. Finally, I formally describe the criterion of meaningfulness that will be used for the construction of the final model.

The final substantive chapter of the dissertation should work as a full explication of the model I am proposing, incorporating the entirety of the philosophical and formal apparatus I have been discussing in the previous chapters. The formal model is a complex structure, comprises three levels of intensional structures that should support the epistemological theory I opted for within the margins of this project. The basic structure that supports the account of verificationism I have proposed will be defined as a *semantic screen* for defining candidacy of statements of any natural language for expressing propositions. Only and all of those statements that pass the screening can be considered meaningful by means of the formal intensional criterion I have established. The binary accessibility relation for the criterion of meaningfulness for this level of the structure is validated by an S5 system, meaning that it is an equivalence relation. This constitutes the first level structure of the model.

The second level structure is the epistemic-doxastic superstructure that is fed statements that passed the test of semantic screen of meaningfulness from the first level. This structure is central to the dissertation, as my first (and for quite a while the only) goal was to show which epistemic and doxastic logics can be naturally adapted for modelling natural inference in epistemic and doxastic terms. As this project was expanded in a way that it warranted formal representation, or rather support, for the theoretical notions that were introduced along the way of constructing this thesis, the remaining intensional levels were added to the model. The epistemic and doxastic level that I introduce in the chapter is by far the most complex, as it validates two distinct accessibility (or membership) relations, one for knowledge, and the other for belief. The accessibility relation for knowledge is validated by the axioms of the logic S4.2 and is supported by a Kripke frame, which makes it a partial preordering relation. The accessibility relation for belief is quite more complex, as it uses a conditionally defined modality, which is originally supported by an ordering plausibility model. Furthermore, the plausibility model is easily adapted to the structure of a neighbourhood frame, which is finally shown to be augmented by the relevant axioms and rules of inference (K and Nec) and explicated through a Kripke frame. This set of transformations (or rather translations) proved that the entirety of the epistemic and doxastic level of the model is normal, which was promised by the very name of this dissertation.

The final level of the model is an intensional superstructure layered on top of the epistemic and doxastic level, which formally captures the notion of algorithmic knowledge. Again, in order for something to be fed into this level of the model, it first ought to pass through the first, and then the second level. All and only those statements that have climbed the ladder through the defined levels of the model can be thought of in terms of algorithmic computation. In this chapter of the dissertation, I will attempt to show why I used the Distributed Systems Models to capture the epistemology of verificationism. We will observe through several examples how a multi-agent systems works by using the introduced formal instruments. The examples of the final chapter are centred around everyday situations and should clearly delineate why I have opted for this specific formal inventory in order to capture the epistemic and doxastic realm of inference. Furthermore, I will display how we can enrich the model (if there is a need) with new epistemic and doxastic operators, such as the operator for common knowledge and another for distributed knowledge, without compromising any of the properties of the system. All that is de facto inferable by the agents of the system can be shown to be computable by the agents when implementing the appropriate algorithm for solving the specific problem. Furthermore, I will again return to the notion of partial algorithms that was introduced in the third chapter, and will show why some algorithms can be used for solving multiple problems.

Finally, I will conclude the book with suggestions for further research. Just to give the reader a brief preview, the first suggestion pertains to the work on defining motivational content for local states of the agents of Distributed Systems. The second suggestion revolves around the notion of combining dynamic doxastic and epistemic logics with STIT systems, in order to devise a potentially useful approach to game theory with pseudo-dynamic systems such as CDL with appropriate properties.

CHAPTER II – MODALITY IN REASONING

CONDITIONALS AND MODALITIES

Beyond the realm of the here and the now, there are statements, or maybe even facts, that deal with how things could have been, might be, or ought to be, and anything in between. Modalities are essentially instruments with which we calibrate our thought and language to deal with the notions of possibility and necessity (Kratzer, 2012). They provide us with dispositions to analyse situations not only in terms of how things actually happen, but in all the *possible* ways in which they might. Let us take an example of an automobile mechanic. When he was studying to become one, there was only so much that he could learn from dealing with a working engine. He could learn to recognise its parts and what does the system look like when everything works as intended. But being a mechanic certainly does not imply dealing with machines that fulfil their function properly. Not until he has become engaged with various faulty ones, along with learning about possible approaches on how to resolve their issues, would he probably say that he *understands* the inner workings of engines. The list of situations of what can go *wrong* with the engine and learning what options exist to repair it appear to carry vital information about what an engine *is* in many ways. This approach to understanding the world around us is not out of the ordinary in the slightest. We tend to analyse our surroundings not only in the ways that they are, but in many ways they could be. It is an essential part of our reasoning and it seems that our interactions with our environments would be greatly hindered if we only had at disposal the things presently are.

To being with, I would like to very briefly observe the motivational history of constructing the formal apparatus for dealing with the problem of modal statements, after which I intend to offer a provisional and non-exhaustive list of various topics related to modalities and modal linguistic categories that we find in our everyday lives in order to get an idea of how it affects our reasoning and understanding of the world around us. This chapter of the dissertation is largely informal and should only serve as a general introduction into some motivations for developing intensional semantics from a natural language perspective. Formal systems that deal with modalities have not been developed until the last one hundred years with Lewis (1918, 1932) starting to shape an axiomatic framework for dealing with unpalatable consequences of understanding the conditionals of natural language as material implications. This was soon proven to be fertile ground for dealing with conditional structures in our everyday discourse.

As it is quite clear by now, modalities are often bound to conditional statements in our everyday reasoning. Or maybe rather, conditional statements are generally translatable into modal ones upon closer inspection. (Stalnaker, 1968) This is quite obvious to anyone who has ever picked up virtually any introduction to contemporary modal systems (Garson, 2006; Chellas, 1980) because of the simple fact that most such texts contain a major chapter on various treatments of conditional statements. Even more so, many logics, among which some will be relevant throughout this dissertation, are based on conditional structures and it can easily be seen how they are related to our pre-theoretic conceptions of modal terms. For instance, the methodology of natural sciences that is based on conditional statements can be understood in two ways; (1) either we observe some state of affairs in the world as initial conditions and attempt to infer what might constitute its consequence, or otherwise (2) we observe a phenomenon as a consequence of some states of affairs and attempt to reconstruct the initial conditions that lead to it. In other words, we rationally construct or reconstruct models of the world around us using some available information. The first has to do with our capacity for *anticipation* of further behaviour of the observed systems, while the latter has to do with our propensity for constructing *explanations* of the systems' behaviours. (Popper, 1959; van Fraasen, 1980). None of those kinds of reasoning would be possible if we hadn't used hypothetical thought. In other words, it appears that conditional reasoning gives us a network of possible situations that resulted in the state of affairs we are currently in, along with ones that we are heading towards. Even further than that, we are able to conceptually observe how the situation would have been different *at this point in time* if things went down a different path. This theoretical framework is usually described as the plane of possible worlds or situations. So, we can say that it provides us with the logical structures for dealing with conditional reasoning.

Canonically, we recognise two main types of conditionals, and although both play an important role in understanding modalities, one is of essential interest to the domain of modal reasoning. The two types are indicative and subjunctive. Indicative conditionals give us information about how things usually work *ceteris paribus* (Edgington, 1995). If the train arrives on time, I will not be late to work. If it starts raining tomorrow, the plants in the garden will not die from heat. When reasoning about these kinds of conditional statements, we usually isolate some perceived variables from our environment that we see relevant for the behaviour of the system and establish relations between them. As we have seen, they serve the purpose of anticipating further developments given some regularities that we have thus far perceived or inferred to about the behaviour of the observed system.

The latter, the subjunctive kind, is particularly interesting for this theoretical framework, as it deals with things beyond their *de facto* perceived behaviour, and ventures into pure relationships between them, i.e. not taking into account their actuality (Stalnaker, 1968). Subjunctive conditionals can be further subcategorised into two types; (1) counterfactuals on one hand, and (2) 'pure' hypothetical subjunctives on the other. The first subcategory is *irrealis in form*. It pertains to the situations which could once have been constructed as viable options, but no longer are, or the ones that were only possible under stipulated conditions that were, for instance, never possible given how the system is arranged. The second form, the ‘pure’ subjunctives still do not pertain in any respect to the actuality of things, i.e. to how things in fact are, but only to the relationship between some possible states of affairs. They are not of *irrealis* form, as their conditions can still be fulfilled. A similar analysis can be found in Bennett (2003).

Actuality can, in this kind of discourse, be viewed as an indexical – in the framework of modal statements, i.e. out of the given possibilities, one of them is the actual one, the one we inhabit. This is not in any way a metaphysical statement, but only conceptual, as it only pertains to the way we engage with reasoning about modalities and hypotheticals.

So, as we have so far discussed, the counterfactual conditionals, as the word itself implies, deal with situations that are outside the scope of actual possibility, i.e. they pertain to not-anymore-viable options of how the world will behave. The examples can be construed as following: “Had I chosen to study sociology instead of philosophy, I wouldn’t be writing this dissertation right now.” The event in the antecedent obviously hadn’t taken place, but it appears that we are capable of addressing it coherently. So it appears that we are capable of determining a possible outcome of something that has never taken place and can never take place, in virtue of our understanding of some regularities about our environment. The reasoning in this case is strictly calibrated to the way we use modalities in our everyday discourse.

And although the subtype of subjunctives that got most philosophers’ attention is the counterfactuals, the ‘pure’ hypothetical subjunctives that are not counterfactual are just as interesting. We can imagine an example in which Jones arrives at the doctor’s office with a set of symptoms. The doctors examines him and says “Had you also had a strong headache as an addition to your current symptoms, the most probable diagnosis would be sinusitis. Otherwise, I believe you have a simple case of common cold”. Unbeknownst to him, Jones might have had a headache that he either forgot about or forgot to tell him about. This structure is definitely not *irrealis* in form, as it still constitutes a viable option, but it behaves more similarly to a counterfactual than to an indicative conditional statement because of its grammatical and logical form.

We can argue that those two kinds of conditionals open up the intensional plane of inference and are capable in one way or another of supporting modal structures of natural language (Stalnaker 1968). However, at this point in the chapter, given what we have observed thus far, it might be proven useful to address come examples of usage of modalities in our everyday discourse. As I have stated, the examples that I introduce do not in any exemplify all of the elements of an exhaustive list of modal categories, but are rather adapted to the needs of this dissertation. The philosophical peculiarities of each (or at least most) of them will, hence, be addressed in further chapters of the dissertation and some will be formally captured by various modal systems that we will observe. They might shed a sliver of light onto how we, in fact, use modal terminology and to which kinds of our reasoning they can be applied. The examples that I will propose below have been chosen in order to be calibrated to the theoretical specificities of the theories that we will observe, but similar examples can be found in the works of the authors I have cited in this chapter, such as Bennett (2003), Edgington (1995), Stalnaker (1968).

MODAL STATEMENTS AND CONTEXTUAL USE – EXAMPLES

In the vein of what we have observed in the discussion on conditionals of natural language, we might notice that use modalities that help us codify information about patterns – "If she were at the office, the door wouldn't be locked." This is an instance of the type of reasoning that we describe as abductive or otherwise as inductive. Abductive reasoning is conceptually bound to the intensional plane, as it is by definition grounded in the *possible* explanations of the perceived phenomena (Biggs, 2011). It pertains to patterns that we recognise in what is often called in philosophy *ceteris paribus* – the unchanged conditions. All of the surrounding circumstances indicate a fact, as the fact usually occurs when they occur. This mode of reasoning is obviously fallible in as much that it does not guarantee salva veritate condition, but it is also the reason we interact with our surroundings so successfully. The systems' success in interacting with its environment often depends on the generalisations that are made about it. If a puma never generalised over the behaviour of its prey, it would certainly never catch it. If we never generalised over the behaviour of the traffic when we go to the office, we wouldn't ever start our cars. Even though many kinds of generalisations may be proven to be dangerous in axiological contexts, it appears there exist many areas in which we wouldn't be able to function without them.

Among such statements, there exist modalities of warning with implicit conditions "You must not venture into these woods – there are wild beasts wandering around." Prima facie, this looks like a prohibition, however, when we take a closer look, it has an implicit condition. What it actually says is "You must not venture into these woods *if you care for your wellbeing*." This is an example of attributing motivational states to implicit conditional statements. Later in the dissertation, specifically in the chapter on the analysis of Distributed Systems Models, it will be shown that there exist philosophical and formal problems in attributing motivational states of agents within the system when dealing with conditional statements (Stalnaker, 1999). I will attempt to show that there is a substantial difference in formally capturing the understanding of mental content when interpreting the motivational states in conditional form, even though it appears to be the most natural reading of them.

A similar modality is used for expressing non-motivational prohibitions; "You cannot be here, this is for organised groups only." At first, understanding plain prohibitions as modalities might appear odd to someone who hasn't thought about it much. However, they can be understood as limiting the number of viable options for agents' actions in order to steer the behaviour of the entire system. This can also be argued to constitute a motivational prohibition, however, only when perceived in a multi-agent system. Similar examples of prohibitory modals can be found in Horty (2000).

As far as epistemically oriented modalities go, there are such that describe impossibilities under present information; "She couldn't have taken the car, her only set of keys is on the table." This basically states that the only way she could have taken the car is if she had also taken the keys. Since we have evidence of her not having taken the keys, the necessary conclusion is implied of her not having taken the car. These kind of statements play a strictly cognitive role in making sense of our environments and will likewise be dissected in detail when we start observing knowledge and beliefs of agents in the chapter on Distributed Systems Models. They can have quite an interesting philosophical reading in the context of possible worlds, as we can analyse the possibility of there existing a world (or rather, a situation) in which we can model contradictory statements. These kind of analyses can be found in works on hyperintensional modal impossibilities of authors such as Jago (2014) and Berto (2019).

Further along the line, we will find a kind of modality directly pertaining to time - some state of affairs will be capable of materialising only at some point; "Only when the semester finishes will you see how much you have fallen behind your sleeping schedule." These kinds of sentences will be addressed by dynamic systems in the chapter on belief revision and epistemic update. These will allow for new information to affect the epistemic and the doxastic states of the agents in the system. They will also be relevant for defining programmable action – the notion of changing agents' algorithmic behaviour in order to fulfil some function of the system (Halpern et al. 1994).

There are also modalities do not strictly pertain to temporal factors, but to some change of state of the system; "Handing in your resignation means there is no turning back." This kind of modalities we use in order to display the viability of potential states of affairs once some event has taken place. They will also pertain to dynamic modal systems, but in a non-explicit manner. Some systems which will be introduced within the dissertation will not be dynamic *strictu sensu* but will be capable of expressing dynamic qualities through relational structures. One of most relevant systems for these kinds of statement will be the Conditional Doxastic Logic system, displayed in the aforementioned chapter. In this example it seems quite intuitive to understand the statement as pertaining to alethically accessible states. We can reframe it to state that the act of handing in the resignation 'closes' some worlds, i.e. they will not be considered as viable henceforth (Negri and Pavlović, 2023). Even though the modality does not appear explicitly in form of a grammatical category of modal verbs, the statement is clearly analysable through possible worlds semantics. In other words, it appears to be an intuitive instance of conditional usage that warrants the introduction of the intensional plane. These kinds of analyses can pertain to an array of situation that we do not perceive as modal at first, but as we examine their meaning more carefully, it becomes clear that such an analysis offer the most natural explanation. Now that we have inspected some relevant examples, it might be appropriate to say a few words on the intricate relationship between modal language and the cognitive processes when acquiring doxastic and epistemic states.

MODALITIES, VALIDITY, AND BELIEF FORMATION

Generally speaking, when dealing with the issue of modalities in reasoning it is essential to understand that they play a vital role in our pretheoretic conceptions of justification and reliance on evidence in reasoning processes. Our protoepistemological conception of evidential reasoning surprisingly works very naturally in parallel to our understanding of logical validity and soundness; when we say that we have conclusive evidence of φ, it can be viewed as an effective translation into "you *must* believe that φ", not dissimilarly to "if it is the case that if the premises are true, the conclusion is *necessarily* true, then the argument is valid." This deontic modality is integrated in the way we perceive the *guarantee* of truth that the conclusive evidence provides us with, while the alethic modality in the ubiquitously accepted definition of logical validity (the conclusion is *necessarily* true) states in parallel that no model is constructible that validates the antecedent while not validating the consequent.

In effect we can perceive both as normative statements about the requirements of rationality. Moreover, even if we were faced with inconclusive evidence of some proposition φ, the formulation would be quite similar – "you *should/are recommended to* believe that φ is the case given what we have learned". The same kind of normativity, albeit weaker, applies to case. This obtains because the modalities that we have discussed thus far do not pertain to the degree of certainty (as would one naturally assume given their role in natural language), but to the *kind* of reasoning we implement when faced with evidential support of a proposition. So when we say "φ simply *must* be the case, because x, y, and z", we are claiming that there are no epistemically viable options to consider. From the set of options that rational agents could observe, and consider viable, the evidence that we are presented with excludes all of the considered options but one. I will attempt to show with the stress on epistemic externalism throughout the dissertation that the conception of rationality and being receptive to evidence has a broader intended domain than human reasoning. We can ascribe epistemic and doxastic states to all kinds of systems that are calibrated to track some variables in their environment. In other words, they are programmable for action in a way that they behave in a certain manner when some proposition obtains, and not behave in such a manner when it does not.

Take for instance a vehicle. When it runs out of oil in the engine, the oil lamp lights up to give you a warning that if continue driving in the present conditions, specifically in this case the absence of oil in the engine, your automobile will break down. In effect, the behaviour of the automobile can easily be framed as it being responsive to evidence. If we might object that it lacks internal conscious states in order to define it as being responsive to external evidence, we might be barking up the wrong philosophical tree. The epistemically externalist approach to modality and evidence solely implies the following analysis: an input variable to the system becomes bound to a set of possible states of the system, and while the system's computational capacity provides an output, the set of possible states collapse into one.

In our example with the oil lamp, the lack of oil is obviously the problem. However, it is not the lack of oil that is the input variable, it is the electrical impulse from the engine to the board computer, just as the lack of water in a human being's body is not itself an input variable – the feeling of thirst is. The impulse in the form of input variable is yet to be interpreted – for instance, it can mean several distinct things. For one, it can mean that the wiring from the engine to the board computer is faulty, second it might mean that the sensor for detecting oil level is faulty, and finally three, it may mean that the engine lacks oil. The contemporary board computers can be very successful in determining which of the three situations was the case by testing it. It performs a sort of a crucial test; it isolates the variable that is different in the three cases and it constructs a test around it. If it cannot construct a test around a variable that is different in all three cases, then it might implement two separate tests if needed where it eliminates one option and proceeds to eliminate the remaining ones until the possibilities collapse into one. Finally, the oil lamp turning on *is* the output of the system, or to the driver, the evidence of the automobile being responsive to evidence.

This electrical impulse in the automobile that worked as our input variable resembles our nervous system when we feel a sudden pain in our leg. The principle is the same, we just tend to frame it differently because of the phenomenological aspect of it. Before we analyse what happened, the pain can be interpreted as the input variable, and our interpretation of it can pertain to different situations that *might* have caused it. For instance, we can think that either our nerve inflammation that is acting up, the cat bit us in our sleep, or we stubbed our toe on the wooden frame of the bed. Once we have analysed the situation, we determine the reasonable course of action. Either we will banish the cat from our bed, or we will put some local analgetic on the affected area, and so on. We exclude the possible sources of pain by means of collecting evidence from our environment (playing the role of additional input variables) that will help us determine which explanation of the original impulse was the most likely. This framework will be developed in more detail in the chapter on Distributed Systems Models, as it will be captured by the notion of a simple processing agent that is programmable for action based on tracking variables in their environment. So, as I have discussed, I will attempt to show that such agents are so unspecific that the intended domain of the model's application extends beyond the usual application for epistemic and doxastic modelling. Just for clarificatory purposes, we might want to observe which kinds of modalities are generally considered by the authors within the field of philosophy of (modal) logic.

In the relevant literature we can find that theoreticians recognise several different kinds of modalities in one framing or another: logical (alethic), metaphysical, nomic, temporal, deontic, dynamic, epistemic, and doxastic (Kment, 2012; Garson, 2006). There might exist some friction within the field about some of them being trivial or reducible to others, but this is the picture in broadest strokes. Alethic modality is usually considered to be the cornerstone for understanding the other kinds, as it does not have an intended domain – its domain is, simply put – the truth of the matter. While that could be said about metaphysical and nomic modalities as well, it appears they are narrower in their scope, as they presuppose some truths that the alethic one didn't. They appear to be more specific towards the ways how *our* world is structured. Temporal and dynamic seem to be intimately connected as well, as both are intended to describe some changes or a lack thereof of some states of affairs. Their structure allows us to handle any sort of progression in an observed system. The deontic modality is the only normative one, providing us with a toolkit to describe what we ought to do or are required to do in order to achieve something, they allow us to speak about permissions and restrictions. Finally, the ones that are the most interesting for this dissertation are the epistemic and the doxastic modalities.

At first, it is not quite clear as to why knowledge and belief are modal concepts at all. We know what we know, and we believe what we believe – what is modal about that? Well, upon closer inspection, when we really think about how we use those terms, it becomes quite apparent that when we talk about what we know or believe, we are excluding some ways that the world might be from our theories of the world based on what is *compatible* with what we claim to know or believe. In as sense, with these modalities we are exploring *viable options* of how the world might be based on what we have learned until now.

Furthermore, as I have stated in the introduction to the dissertation, the two concepts arrived in the same philosophical package some odd 2000 years ago, and we would be far off the mark to we claim convergence on the topic of how they are exactly connected. For the better part of those two millennia, the three-part definition of knowledge – knowledge as justified, true belief – was famously rejected by Socrates in the Plato's work, Theaetetus, but appeared to have convinced everyone else and served as a basis for developing epistemological theories until the 20^{th} century. In the 20^{th} century, Edmund Gettier (1963) wrote a very short paper about a problem in the three-part definition that was demonstrably recognised and understood before he wrote the paper, but has somehow fallen into obscurity of collective philosophical amnesia. Gettier's paper dealt with the problem of establishing criteria for justification for some contentious epistemic situations, and his work propelled the field of epistemology into working out the ways in which this criterion can be rectified in order to conform to our pretheoretical conceptions of knowledge.

Offering a theory of knowledge and belief might just be a more complicated issue than it probably initially appeared, and in order to get some clarity in developing such theories, philosophers have adapted formal systems to be able to clearly and succinctly define the abstract interrelationships between them. The formalisations themselves may have appeared to have caused some unwanted idealisations when explicating the concepts of knowledge and belief in formal settings, however it soon became clear that this was not a symptom of the formalism, but of philosophy in general.

In this dissertation I will attempt to construct an epistemic model that will be based on epistemic and doxastic modalities and show how its inventories are useful to depict epistemic situations that are present in our everyday lives. Furthermore, I will attempt to show how a formalism can support a philosophical theory in a way that the statements in the formal setting reflect the statements within the theory.

CHAPTER III – EPISTEMIC AND DOXASTIC LOGICS

"We begin by modelling phenomena like knowledge, belief, and desire using mathematical machinery, just as a biologist might model the fluctuations of a pair of competing populations, or a physicist might model the turbulence of a fluid passing through a small aperture. Then, we explore, discover, and justify the laws governing those phenomena, using the precision that mathematical machinery affords."

(Pettigrew, Weisberg, 2019, p. v)

EPISTEMIC LOGIC AND RESOURCE-BOUND AGENTS

The construction of logical systems that encompass the notions of knowledge, belief and subjective indistinguishability has historically been shown to be quite an arduous task. The original proposal of developing a formalism which would allow us to talk about these concepts in the most abstract of ways was introduced by Hintikka in 1962. His proposal set its roots in modal logic, introducing intensional semantics into the enterprise of what was soon to become known as the first epistemic modal logical system. The modal operator, □ ('box') was epistemically and doxastically interpreted to K and B operators and were then understood as formal representations of knowledge and belief as the intended domain of the relevant modality shifted from alethic to epistemic and doxastic one. Furthermore, this system also introduced the notion of agency, as it would seem somewhat unreasonable to speak about knowledge, belief, and similar concepts without introducing a bearer of them, i.e., without having someone or something that we can ascribe them to.

This was generally regarded as a successful programme, and was, hence, further developed by Hintikka and his peers in order to remedy some unpalatable consequences of his proposed system. One of such consequences was the problem of logical omniscience. In short, the problem of logical omniscience follows from the deductive closure of the system in question, warranting that the knower s of some proposition φ is, by virtue of knowing φ, invariably in the state of knowing all of the logical consequences of φ. Or, in doxastic terms, if he believes that some proposition φ obtains, by virtue of believing φ, he is invariably in the state of believing all of the logical consequences of φ.

This is obviously not the case, as we can easily imagine a plethora of counterexamples. For instance, we might easily be in the state of knowing Peano's axioms, but not quite as easily be in the state of knowing all the theorems of the classical arithmetic. Conversely, we might be keen to accept that Syd Barrett was the chief composer of Pink Floyd's debut, but would not say that we accept that Roger Keith Barrett authored the majority of the album Piper at the Gates of Dawn, although these two statements are, in fact, materially and logically equivalent. This is an obvious consequence of dealing with resource-bound, non-omniscient agents. In other words, a system that contains deductive closure is calibrated for another type of agent – an idealised one. An obvious fact about this problem is that it is not a problem for the system that validates such inferences, but the problem of divergence between an inferential apparatus that we endorse and the intended domain to which it is applied. Logical omniscience, as a consequence of the system causes no problems for the system itself, as it does not generate any contradictions nor hinder the behaviour of the system's inferential inventory. It is solely a matter of applying an idealised logic onto a non-ideal agent who should use it or be described using it.

The problem of logical omniscience can formally be explicated in several ways, as it can pertain to various types of deductive closure. The most commonly discussed one is of the form:

DEFINITION. $K(\varphi\rightarrow\psi)\rightarrow(K\varphi\rightarrow K\psi)$.

This type of closure is usually intrasystematically defined as an axiom – the K axiom, or the Axiom of Distribution. It states that if some implication is known, then if the antecendent of the implication is known, so is the consequent. Upon close inspection, it is a direct consequence of epistemic closure under material implication. This means that for every instance in which the epistemic modal operator is distributed over the connective of implication, this type of omniscience obtains. Not only is it present in all the normal modal systems, but it constitutes one of the two conditions for normality of modal systems. Generally speaking, normality can be viewed as a desirable property of the system, as it guarantees that the system is adequate, complete, and decidable.[1] All of these properties indicate the deductive strength of the system, which is why we are keen on maintaining them in the models of knowledge and belief.

Apart from the Axiom of Distribution, the other condition for normality is the acceptance of the rule of inference entitled Necessitation. The closure under the rule of Necessitation is the second most common form of logical omniscience in present day literature and is usually defined as following in the epistemic modal systems:

DEFINITION If $\vdash \varphi$, then $\vdash K\varphi$.

[1] The proofs for this being the case will be contained within the penultimate chapter of the dissertation on modelling verificationism. Even though the proofs will not address the entire classes of logic, I will show for all the relevant systems for the development of this model that they possess the said properties.

This form implies that every agent knows every theorem of the system. Seeing this is not the case, it can be viewed either as an idealisation that the system accepts in order to maintain some of the desired properties, or it can be viewed as a problem that is in need of solving. The attempts to resolve the problem of logical omniscience are usually described in terms of *hyperintensional systems* – establishing a finer-grained model on (usually) syntactic basis that provides a possible solution to this problem. This is not an essential part of my endeavour of constructing an epistemic-doxastic model, but a few words about such systems should be said nonetheless, as they constitute such a major part of the current discussion on modelling knowledge and belief.

HYPERINTENSIONAL APPROACHES

Hyperintensionality is generally based on a rather simple idea of constructing systems which are more fine-grained than the classical normal systems of modal logic. The authors who are renowned to implement hyperintensional approaches are Mark Jago (2014), Francesco Berto (2018), Eric Pacuit (2017), Graham Priest (2017), Joseph Halpern (2007), Kit Fine (2015), etc. The fine-grainedness is usually captured by some of the following approaches, which I will briefly describe in order for the reader to get an idea how my attempt of constructing a model differs from the rest:

(a) The syntactic approach, which most commonly takes sentences instead of propositions to be the constitutive elements of the system, making the syntactic formulation of the sentences a relevant aspect of discrimination between the possible worlds in which they obtain or not. This seems like an appropriate solution to some formulations of logical omniscience such as the epistemic closure under material implication. In this case, if an agent knows *p* while there being the case that *p* implies *q*, she still might not known the *proposition q*, as the sentence that contains the proposition might be syntactically distinct from the one that serves as a consequent of the implication. This approach is discussed quite clearly and in great detail (but not defended) by Robert Stalnaker in the paper The Problem of Logical Omniscience I (1999), in which he frames the issue with the instrument he entitled 'the belief box'. He shows how the authors such as Jon Doyle (1979) attempt to establish a model on the syntactical basis of natural language, where beliefs are stored within an abstract structure in the mind of agents, the belief box, as sentences that express doxastic attitudes of agents about the world. Stalnaker comments as the approach appears to rid the system of the problem of logical omniscience, but also shows the pitfalls of such an approach with *knowledge availability* being the central issue. He shows that if we fail to offer an account of this kind of availability, the syntactic approach falls short of a full explanation because the knowledge might be there for an agent, but she has no way of getting to it, let alone using it.

(b) The second prominent approach to solving the problem of logical omniscience in hyperintensional frameworks pertains to modelling impossible worlds as candidates for being considered viable options for our notion of knowledge. This virtually means that the logical system that we opt for is capable of maintaining contradictory states of affairs in the worlds such that $w \in W$. This approach allows for epistemic closure under material implication because it models the possibility of φ being inferable within the system, in this case the distributed consequence of the Axiom of Distribution that states that the agent knows the relevant proposition, while being the case that the agent does not actually know it. This kind of approach tolerates this contradiction by using paraconsistent and paracomplete systems that validate it. One such endeavour is discussed by Berto and Jago (2022) in the monograph *Impossible Worlds*. Generally speaking, this approach is regarded as quite contentious by an author keen on maintaining the classical account of propositional logic, as the system that are paraconsistent and paracomplete seem to lack any natural application on the epistemic and doxastic situations that we want to model. Hence, the consequence of using this approach to resolving the problem of logical omniscience is relinquishing the postulates of the classical account of propositional logic.

(c) The third and final hyperintensional approach to resolving the problem of logical omniscience is introducing an *awareness function*. The awareness function can be established in several distinct ways; either as a modal operator which will state that the agent is or is not aware of some of the proposition that she knows or believes, or as a function from the set of epistemically accessible states into the set of states that the agent is aware of. Obviously, the function is neither injective nor surjective, as the function is not such that its image is the entire codomain. In other words, there exist such elements of the domain that are not mapped onto any element of the codomain, as the agent is not necessarily aware of each proposition that she knows or believes. This allows the system to tolerate the instances in which the forms that we stated to represent logical omniscience are valid, while still maintaining that the agent might not be aware of their consequences even though they hold within the system. In this dissertation I will attempt to defend something akin to the awareness approach. The apparatus similar to the one that I opted for was discussed in detail by Halpern, Moses, and Vardi in their paper Algorithmic Knowledge (1970), and will be covered in the next chapter of the dissertation. This being the case, I still do not intend to introduce an awareness function as an instrument of the base logic of knowledge and belief, which will remain normal, but only as a superstructure that will determine what knowledge is computable for an agent in certain epistemic and doxastic situations. The superstructure that I introduce will not affect the logic of knowledge and belief in any respect, as its axiomatic schemas and inferential apparatus remain unchanged. Furthermore, although my account will validate such a superstructure, I will use a slightly different terminology – the function will be simply defined as an algorithmic knowledge function, which will be described and explicated in detail, both philosophically and formally. This superstructure will, in fact, work as a solution for logical omniscience for knowledge, with the base theory of knowledge and belief remaining unhindered by formally weakening them. As opposed to knowledge, I will deal with the doxastic part of the model in a different fashion by dynamising belief through a conditional logic with full pre-ordering relation. As I have stated in the introductory chapter of the dissertation, this logic of belief will be supported by a plausibility model, which will formally integrate, but philosophically weaken the problem of logical omniscience.

Now that I have briefly described the main approaches to resolving discussed problem, I will turn back to Stalnaker's analysis that will help us flesh out some relevant properties of systems that integrate logical omniscience into their inferential base.

In this part of the dissertation, my intention is to address epistemic and doxastic modelling through the lens of Distributed Systems Models with a focus on logical omniscience. I will discuss further in the dissertation that Stalnaker's take on this issue diverges quite radically from the attempts of dealing with logical omniscience made by authors such as Fagin, Halpern, Moses, and Vardi (1995), and Jago and Berto (2014). Such authors are often more inclined to weaken either the logical system in question or its application through epistemic modelling so the system would pertain to a broader set of agents, specifically not necessarily ideal ones. Stalnaker, diversely, sees this idealisation as a necessary course of action in order not to compromise the deductive strength of the system in question. Hence, the system and its application that he proposes will still suffer logical omniscience as a symptom of keeping the logic as close to the classical account as it can be because of its desired properties.

As we have observed, the majority of other relevant authors in the field of formal epistemology, conversely, attempt to weaken the Problem of Logical Omniscience by means of introducing hyperintensional models which facilitate the possibility of differentiating propositions based the syntactic properties of the sentences that express them. These attempts of weakening the Problem of Logical Omniscience fall outside of the scope of this dissertation, but certainly play a prominent role in pulling the proposed epistemic model closer to resource-bound agents such as ourselves. The logic that such models validate will obviously be deductively weaker in some respects, as some principles of closure will have to be abandoned. However, as I have already stated, we will not venture into these endeavours, as the models (and their agents) that hold interest for this dissertation are, in fact, idealised.

The prerequisites for agents to be described as idealised are that (1) they have infinite storage capacity (for whatever type of information the system is calibrated to), and that (2) they do not require time for processing the inference, rather for it to be instantaneous, as the system is neither strictly inferential, nor dynamic, but a mere set of relations between abstract objects. Evidently, no finite physical system is capable of attaining this standard. In fact, relatively speaking, even the most powerful supercomputers with enormous data storage and extremely fast computing abilities are infinitely closer to a human being in terms of their processing power than they are to an idealised agent, which such a system would require. So why would anyone even attempt to introduce such systems which are so far removed from the real-world setting that we are trying to capture? And is there an alternative approach to this problem through epistemological theory?

The following part of this chapter is of expository character to some extent, basically explicating the required apparatus for knowledge and belief ascription within a formal epistemic model. Conversely, the second part will be focused on shedding some light on handling logical omniscience within a formal epistemic theory, along with establishing the prospects of elucidating information accessibility for ideal, and resource-bound agents within the framework of Distributed Systems. I intend to consider some attempts of clarifying these issues by Stalnaker (1996, 2001, 2006), Fagin, Halpern, Moses, and Vardi (1970, 1995), comb through their analysis of this issue, and finally observe how we can go about either weakening or trivialising logical omniscience when constructing an epistemic and doxastic model. In short, the following is the general outline of this chapter.

First, I will briefly discuss the relationship between a formal system and the epistemic theory that should be supported by it. I will attempt to clearly display the infrastructure of the model and will show how its elements interact. Secondly, I will introduce three modal systems as a sort of a case study, upon which we can observe how manipulating axioms generates various restrictions on accessibility relations, and thus changes the metalogical properties of the systems at hand. The three logical systems that will be taken as case studies are S4.2, advocated by Stalnaker (2019) and Lenzen (1978) as the 'correct' logic of knowledge, S5, the strongest modal system with a universal accessibility relation, and KD45, the most common logic of belief in static modal settings. After that, I will introduce the frame of Distributed Systems Models in a static epistemic and doxastic environment. This framing of abstract situations pertaining to knowledge and belief is incredibly useful when dealing with multi-modal and multi-agent environments, and I will attempt to show how its application can be useful for the model that is the end-product of this dissertation, establishing a very natural reading of the account of verificationism that I propose in the dissertation.

THE RELATIONSHIP BETWEEN FORMAL SYSTEMS AND EPISTEMOLOGICAL THEORIES

The focal point of understanding the basis, as well as the pitfalls, of epistemic modelling is grasping the need for a sort of complementarity between the formal infrastructure and the theoretical framework of the model. The model itself might be nothing more than a set of axioms and theorems which they entail, but its theoretical interpretation just might give us a clue whether the model is doing what it is supposed to – modelling the behaviour of its intended agents and our respective ascriptions of knowledge and belief to them. We are in the position to ascertain whether our previous theoretical conceptions of knowledge and belief could have any merit in constructing a clear, consistent, and unambiguous picture of the most general of epistemic situations. We would certainly hope they would. However, this might just not be the case at all.

Firstly, it needs to become clear that the epistemic theory of our logic of knowledge and belief will comprise few variables. This is primarily a pragmatic decision, as if it were otherwise, the model would become quite intricate, and by default, deductively weaker. This virtually means that we ought to carefully choose which theoretical concepts we want to ascribe to their formal counterparts. We want to define a system that is sufficiently expressive such that, at least to some extent, captures our semantic intuitions of the relevant concepts, but that it is still deductively strong. For instance, we can opt for constructing a combined logic of knowledge and belief with appropriate properties that will correspond to some extra-systematic notions from the domain of epistemology. Combined logics are evidently more expressive, but we need to keep in mind that we must clearly and unequivocally define the technical terms and display their relationship overtly within the logic. This will be covered in some detail in the chapter three on restrictions on accessibility relations. There we will also try to examine how to maintain the deductive potency of the system whilst not abiding to the aforementioned principle of theoretical parsimony.

So, what we need in order to be successful in the endeavour of constructing an appropriate model for knowledge and belief are basically (a) a consistent and informative conception of knowledge and belief, and (b) a system of presuppositions, i.e., axioms, which comprise the formal infrastructure of the system, generating some desired properties of the system in question. The first requirement (a) is of interest to the philosophical part of our proposal, while the second (b) is of interest to the logical part of it. These two aspects of our theory are required to be congruent, as the lack thereof can result in a theory that generates consequences that we are not eager to accept. This, in fact, means that when dealing with the philosophical aspect of the equation, we ought to be aware of its implications to the logical part, and vice versa. As we have seen, a well-established epistemic model ought to display an appropriate relationship between the two.

Stalnaker (1999, p.255) points out in his article "The Problem of Logical Omniscience II" that the original proposal by Hintikka, which adapted the alethic modal logic to the domain of epistemology, generates the problem of logical omniscience for non-ideal agents. Since we, in principle, want our theory to pertain to our everyday knowers, this poses a significant issue. If the logic we choose for our theory cannot be properly applied to common, non-ideal agents, at first glance, we must either weaken the logic, so that the domain of its application can become the said non-ideal agents, or we can accept that the model is only pertinent to ideal knowers. But there is also a third option. It is somewhat less common in the present discourse, but is still a valid way to try to resolve or, at least weaken the Problem of Logical Omniscience. It constitutes the revision of the concepts of knowledge or belief, so that the system can maintain its desired properties, while becoming applicable to the common, non-ideal knower.

But how exactly do we explicate these conceptions of knowledge in a formal modal setting? This is a question that requires us to talk about the semantic notion of possible worlds. Semantics of modal logics are intensional, which means we need to show which worlds are epistemically or doxastically available to our knowers. The frame F of our system will, hence, comprise such a set of possible worlds W, but also the relation of accessibility R_k between them. We will need to overtly display this relation for it to become clear how the concepts of knowledge or belief behave within the model. In other words, we need to show which worlds "see" which, with respect to belief or knowledge of our agents.

Let us look at an example;

Suppose that Jane is in the park walking her dog. She knows that her dog is fiercely afraid of cats, but is perfectly fine around other dogs. She sits on a bench with her dog by her side. She suddenly hears some commotion in the bushes. Suppose that she knows that it either might be another dog, or, God forbid, a devious cat which will frighten her dog immensely. These two situations, or possible worlds, are indiscernible from one another to her. They both pose a doxastic possibility. Now she has to make a decision based on the set of propositions she knows, whether to quickly leave the bench in case of the cat being the culprit of the commotion in the bushes, or to remain seated in case another dog desires the company of her dog. The moment when the said culprit is revealed, Jane's considered possibilities collapse into one. The epistemically possible world that was determined by her seeing either a cat or a dog jump out of the bushes can be understood as the actual world. Within such a framework we can say that Jane did not *know* which possible world she was in before the reveal. So, knowledge can, in this sense, be viewed as pinning down the actual world.

Considering the formal aspect of this enterprise, this example outlines the basic idea of how intensional semantics can be applied to our understanding of epistemic concepts, although the ascription of knowledge to our agents is done in the most general epistemic setting and can be described formally without considering the details of the situation. We need not to think about Jane or her dog, as their particular situation is only an instance of a more abstract formal epistemic picture that can be applied to virtually any situation of knowledge or belief ascription. Such abstract pictures are the direct product of formal epistemic logics. But the logical frame that we introduced is not enough to capture the philosophical implications of these ascriptions, as we so often rely on the concept of truth, especially when dealing with knowledge. The set of all worlds *W* and the binary accessibility relations R_k and R_b tell us nothing of the truth of propositions that constitute a world. They merely establish relations between formal structures, devoid of any semantic content. This is why another element is added, the valuation function $[\![\ ]\!]$. This function attributes truth to a proposition φ in some world that is either epistemically or doxastically accessible to a world in which Kφ or Bφ obtains, respectively. So, the expression $K_s\varphi$ states that the subject s knows that φ, and it is precisely the proposition φ that determines the subset of possible worlds in which φ obtains. Such subset is viewed as the subset of epistemically possible worlds. Mutatis mutandis, the expression $B_s\varphi$ states that s believes that φ, and it is the proposition φ that determines the subset of possible worlds compatible with what s believes, i.e., the set of doxastically accessible worlds. With the added valuation function, we now have a complete infrastructure of the logic with both semantic and syntactic aspect accounted for – a model M. We can, thus, define it as following:

DEFINITION. $M = \{W, R_k, R_b, [\![\ ]\!]\}$.

Keeping this in mind, it still remains to be seen how this abstract system is related to the epistemological theory that we adopted. As we have seen, they must converge in a manner. In other words, knowledge and belief ascription in this formal setting ought to, at least to some extent, mirror knowledge and belief ascription that we encounter in an everyday setting.

Moreover, epistemological theories tend to adopt concepts that can be formally challenging to introduce such as justification of belief. Such concepts could potentially either greatly hinder the deductive strength of a system in question, as they are themselves rather complex and philosophically problematic, or their formal counterparts could become so far removed from their original conceptions, that their introduction itself becomes trivial. So, in order to capture them in a formal framework, one must reconcile some of their properties with what the system is calibrated to do. For instance, epistemic concepts such as justification *can* indeed be overtly explicated in a formal setting, but they will most often not interact with other operators in the system in a manner that would coincide with our theoretical assumptions regarding it. But the formalization of such concepts is beyond the scope of this chapter. For the time being, we’re more interested in examining the relationship between the basic notions of knowledge and belief in a logic and how their respective definitions can affect the problem of logical omniscience.

EPISTEMIC AND DOXASTIC LOGICS AND LOGICAL OMNISCIENCE

As Stalnaker (2006, p.179) discusses in his paper On Logics of Knowledge and Belief, it is quite important to understand which *kind* of agents the model pertains to. We have seen in the introductory chapter that if we were to construct a model based on an epistemic theoretical account that more or less conforms to ordinary knowers, we would need to adapt the formal system or revisit our theoretical account of knowledge and belief in order to avoid full logical omniscience of our agents. To establish this relationship between knowledge, belief, and subjective indistinguishability regarding possible worlds, I will attempt to sketch out and compare two distinct epistemic accounts of logical systems S5 and S4.2 as a sort of a case study. Both these systems are in the class of normal epistemic logics, as they meet the requirements of (1) maintaining all of the tautologies of the classical propositional logic, (2) validating the distribution axiom K, and (3) validating the rule of Necessitation.

Although they are both normal in this sense, they each generate a theoretically biased interpretations of their epistemic modal operators. This, in fact, means that their respective notions of knowledge and belief ascription differ because of their formal inventory. In order to display these notions explicitly, we should examine the relationship between axioms that each system adopts and the relation R from the model which will diverge in the context of the two systems. The relation R is defined through certain formal restrictions that, in fact, model our representations of knowledge and belief within the system. Furthermore, each axiom that is added to the minimal normal epistemic logic will affect this restriction.

It further ought to be noted that two different sets of axioms can define the same logic, based on how they define the accessibility relation. This virtually means that we are offering two distinct formal descriptions of the same inferential system. Let us take two systems, S_1 and S_2 for which such a relationship holds. First we establish that they describe the same logic based on how the binary R-accessibility relation is defined. For example, let the accessibility relation be reflexive and transitive. Now let us presuppose that they differ in two axioms. The upshot is that the axioms of S_1 that are not assumed in S_2 will show up within the system just as well, but as theorems. The same principle holds for S_2 in relation to S_1 mutatis mutandis. For these kind of cases we say that S_1 emulates S_2 iff we use the formal inventory of the former to express the axioms of the latter as theorems.

The following is the list of axioms of S4.2 and S5[2]:

S4.2		S5	
Ax1	All instances of tautologies of CPL	Ax1	All instances of tautologies of CPL
Ax2	Ksφ ∧ Ks(φ → ψ) → Ksψ	Ax2	Ksφ ∧ Ks(φ → ψ) → Ksψ
Ax3	Ksφ → φ	Ax3	Ksφ → φ
Ax4	Ksφ → KKsφ	Ax4	Ksφ → KKsφ
Ax4.2	¬Ks¬Ksφ→Ks¬Ks¬φ	Ax5	¬Ksφ → Ks¬Ksφ

Each set of axioms generates specific restrictions on accessibility relations. These restrictions are definite for each system, but are not necessarily unique. It is possible that two systems with diverging axiomatic schemes produce the same restrictions on the accessibility relations. If this is the case, then it is also the case that both logics have the same set of theorems which are provable within them. In other words, even if two systems have different axiomatic structure, they can still be deductively equivalent. But this is not the case for S5 and S4.2.

S5 is generally understood as the strongest epistemic logic with so called equivalence relation. That basically means that each world in W sees each world, including itself. The universal accessibility relation can be displayed as a conjunction of three weaker conditions on R; (1) *reflexivity*, (2) *transitivity*, and (3) *Euclideanness*. This is a pure epistemic logic, as it is impossible to define another structured accessibility relation which would be stronger that this one. This means that belief collapses into knowledge in S5, since it is not definable in principii.

On the other hand, S4.2 is a normal epistemic logic weaker the S5 but stronger than S4. It was first introduced by Lenzen in 1978, and is currently defended by Stalnaker himself. Its R accessibility relation is also (1) reflexive, (2) transitive, but not (3) Euclidean, as was the case with S5. Instead, it is strongly convergent (see also *strong directedness* in [Chalki et al., (2018)]), which is provable from AX4.2 and the rest.[3]

This means, as opposed to S5, that the R relation is not closed under Euclideanness, which means that belief does not collapse into knowledge, i.e. they do not formally coincide. Such a formal setup provides us with the tools to express both belief and knowledge independently within the system, which means the formal theory is calibrated to an

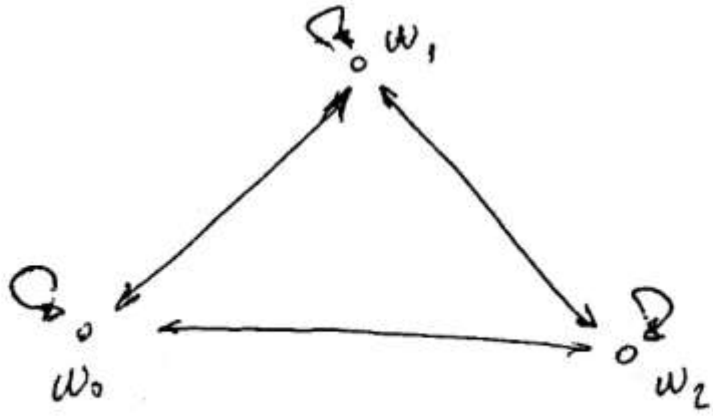
[2] Both systems also validate two rules of inference, MPP and necessitation.

[3] The proof can be found in Tiomkin and Kaminski Theorem 4.11 – soundness and completeness of S4.2 for transitive and reflexive frames with a non-empty final cluster – FC_f = {v ∈ W | (∀w ∈ W) wRv}

epistemology that warrants their distinction. From a formal perspective, this means that the set of worlds that are R_k-accessible does not have the same extension to the set of worlds that are R_b accessible. Thus we are able to formally define this second binary relation R_b for belief operator that does not coincide with R_k. This expansion of the syntactic infrastructure of the formal theory can account for the cases of agents having false beliefs, while still being *logically* omniscient with respect to its knowledge and beliefs. This is precisely why this logic makes for a good case study. Firstly, it indeed can be defined as a combined logic of knowledge and belief. Its pure doxastic counterpart, logic KD45, generates the same theorems as S4.2, with the operator B being substituted for a derived operator of epistemic possibility (¬K¬) (Stalnaker, 2006, p. 184). The Ax4.2 is thus interpreted (by means of understanding belief as epistemic possibility) as the axiom of consistent belief.

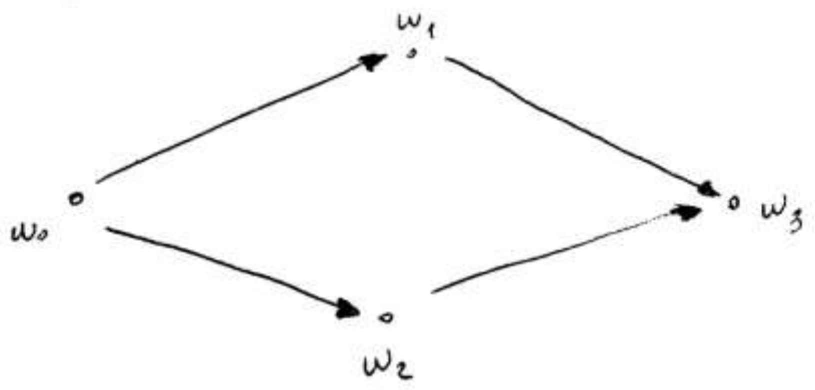
The following is a schematic interpretation of the restriction of the accessibility relation for the logic S5. Each world sees each world in the domain.

SCHEMATIC INTERPRETATION. (ILLUSTRATION.1)



The schematic interpretation for the accessibility relation of the S4.2 system would then look as depicted below. All the worlds that are accessible from the starting world w converge in a single world.

SCHEMATIC INTERPRETATION. (ILLUSTRATION.2)



Now it would be only appropriate to display the system KD45 in the same manner that we did with S4.2 and S5, just to show how it would behave if we were to develop a combined logic of knowledge and belief by constructing a model that validates S4.2 for knowledge and KD45 for belief on one hand, and a model that validates S5 for knowledge and KD45 for belief on the other. The doxastic operator in KD45 will syntactically behave in the same manner as the epistemic operator in the logics that we have seen so far, as it is supported by the same structure, viz. Kripke frames. This means that the binary doxastic accessibility relation R_b that is the result of introducing the axioms of KD45 is also normal in the same sense as the system validates all the necessary requirements for normality – the axiom of distribution and the rule of necessitation. This might be a point of contention for some theoreticians, as it implies that the agents believe all the logical and mathematical tautologies, but as we have discussed so far, it is an unavoidable consequence of attempting to model an epistemic theory with normal systems. Following are the axioms of KD45, which will generate a new specific accessibility relation:

DEFINITION.

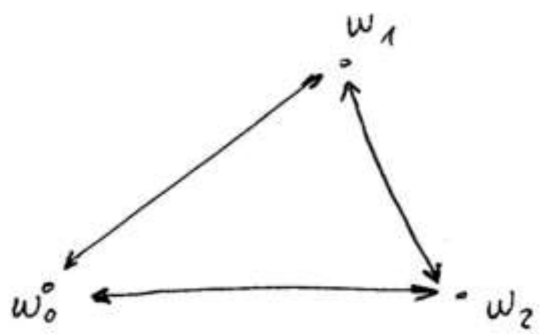
KD45
Ax1. All tautologies of classical propositional logic
AxK. $B(\phi\rightarrow\psi)\rightarrow(B\phi\rightarrow B\psi)$
AxD. $B\phi\rightarrow\Diamond\phi$, equivalently $\neg B\bot$
Ax4. $B\phi\rightarrow BB\phi$
Ax5. $\neg B\phi\rightarrow B\neg B\phi$

As we can observe, the system KD45 is axiomatically very similar to S4.2 and S5 with adapted intended domain for doxastic states, captured by the operator B. The axiom D is also validated in form of a theorem in both S4.2 and KD45. The only axiom that is missing is the axiom of factivity, the axiom T. It is obviously not included as we want our beliefs to be fallible, otherwise we wouldn't be able to differentiate between belief and knowledge. The axiom D that was introduced here did not have to be explicitly included in S4.2 and S5 logics, as it generates the restriction of *seriality* on the binary accessibility relation. Seriality is the minimal expansion of the basic K system and basically means that each world sees at least some world. In order to understand how this restriction is achieved, it might be useful to take a look at its alethic iteration. It is usually defined as such: $\Box\varphi \rightarrow \Diamond\varphi$. As $\Box\varphi$ does not open any new worlds the formula would be satisfiable even if no accessible world is available. As $\Diamond\varphi$ opens a new world, this axiom secures that each situation has access to available options within the model. The axiom D here is defined in a hybrid modal system between doxastic and alethic terms, resulting in the statement that everything that is believed is alethically possible. So, in a way it semantically corresponds to the T axiom for knowledge, where the T axiom establishes relation between alethic and epistemic modalities by claiming that each known proposition is eo ipso true, while the D axiom simply states that each believed statement corresponds to a possible state of affairs in the world.

The frame is then *serial*, *transitive*, and *Euclidean*, which means that the restriction of symmetry is also provable for the system. From this fact, we can see that KD45 is closer to and S5 relation than it is to an S4.2 relation, because it is almost maximally inflated as only reflexivity as a condition on R is not satisfiable in its model.

The following is a schematic interpretation of the restriction of the accessibility relation for the logic KD45:

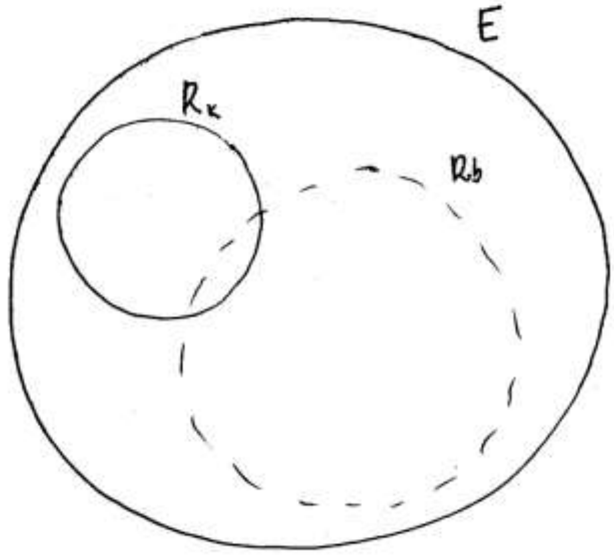
SCHEMATIC INTERPRETATION. (ILLUSTRATION.3)



Now it would seem appropriate to show why constructing a formal combined theory of knowledge and belief when using S5 for knowledge and KD45 for belief would not be appropriate if we want to make our system calibrated for differentiating between knowledge and belief. Every combined model of epistemic and doxastic states ought to have a defined bridge axiom (or rule of inference in rare cases) which should explicate belief through knowledge and vice versa. If we were to establish a model using the proposed systems and their respective restrictions on accessibility relations, we would be justified to infer that that each world which is R_b accessible is also R_k accessible. This is because the R_k accessibility is maximal – an equivalence relation, which means that however we would want to axiomatically inflate the R_b relation, it would still be insufficient. In other words, within such a model, more worlds would be accessible to each world with respect to knowledge than to belief. Philosophically this makes no sense, as we want our belief set to be broader than our knowledge set in order for our model to have the capacity to model false belief of agents within the system.

This might seem counterintuitive at first as we try to define this set theoretically because knowledge implies belief, so we might be keen on asking why would there be elements within the set that is defined by the R_k set that are not contained in the R_b set. Upon closer inspection, it is quite easy to see that this principle holds precisely because the R_k set and R_b set are not populated with propositions that are known or believed, respectively. They are populated by *worlds* that are epistemically or doxastically accessible. The first thing we ought to observe is that the *actual* world is obviously contained in the set of epistemically accessible worlds. This is the case because we have defined knowledge as factive through the implementation of the T axiom. Furthermore, if we were to state that there exist worlds that are doxastically accessible, but not epistemically accessible, that would simply refer to situations in which we have some false beliefs. Furthermore, if the set of epistemically accessible worlds and doxastically accessible worlds were to coincide, then we would be justified to state that the agent in question possesses only true beliefs. Conversely, if we observed a situation in which there is no overlap between the set of epistemically and doxastically accessible worlds, then we would say that the agent in question possesses only false beliefs. Nota bene, we have already seen that since the model validates necessitation, the agents are not capable of having false beliefs about mathematical and logical truths. I have already commented that this is quite a high degree idealisation, but it is a price to pay in order for our model to maintain normality, which results in a plethora of formal properties of the system that we want to keep.

SCHEMATIC INTERPRETATION. (ILLUSTRATION.4)



Now that we have established why the systems S5 and KD45 would not be a good match for a combined model of knowledge and belief as their respective sets of epistemic and doxastic accessibility would be conflated, we should see how would a combined model of knowledge and belief work if we were to take an accessibility relation for knowledge that is weaker than S5 in order to be able to base our formal account of belief on the doxastic iteration of KD45.

Furthermore, because the accessibility relation of the system S4.2 is not defined in terms of the universal accessibility relation, we can construct a more stringent doxastic accessibility relation which will be defined through the defined relation R_k as $xR_by \leftrightarrow (z)(xR_kz \leftrightarrow yR_kz)$ (Stalnaker 2006, pp. 180-183). Also, we can introduce a third relation E, i.e., subjective indistinguishability relation, which can be interpreted as a superset of R_b and R_k relations. Philosophically speaking, that basically means that two possible world x and y are indistinguishable to the agent s in terms of their respective sets of knowledge and beliefs. The relation E can, hence, be defined as $xEy \leftrightarrow (z)xR_bz \leftrightarrow yR_bz$.

Let us for a moment return to the earlier example with Jane and her dog. When she heard the commotion in the bushes, the world x in which the cat was responsible for the noise, and the world y in which it was the dog, are subjectively indistinguishable to her with regards to her epistemic and doxastic states. In other words, both words, x and y are compatible with what she knows and believes, respectively. Imagine that it turned out that it was a dog after all. Now, she might have had a belief that it is the cat, which would, in turn, mean that the *actual* world was in her R_k set – the set of worlds that are compatible with her knowledge, but not in her R_b set – a set of worlds that were compatible with her beliefs. The newly defined binary doxastic accessibility relation R_b within this epistemic model allows her to see the world in which the cat made the noise in the bushes for this epistemic model. However, if she was right about her belief that it was a dog, then R and R_b would coincide, while still maintaining their relationship to E – being subsets of it. This is as far as this example goes for clearing up the intuitions for understanding R and R_b restrictions, but it manages to grasp the general gist of the knowledge-belief interdefinability.

As we have seen in this highly idealized and indeed very simple example, Jane's epistemic and doxastic possibilities, i.e. possible worlds that are compatible with what she knows or believes, respectively, seem quite limited. However, this is not the case. If Jane were to be portrayed as an idealised, non-resource-bound agent, there is plenty more that she ought to believe and know. Precisely, there are infinitely many more things she ought to know and believe as this kind of an agent. She would have to be able to access all the valid conclusions of her belief (R_b) and knowledge (R_k) sets. For instance, she would be warranted to know that "Either a cat or a dog is making the commotion in the bushes or Andromeda is the wife of Perseus." As disjunction introduction is a valid form of inference in classical propositional logic, and she is an agent calibrated to knowing mathematical and logical truths, this would be a consequence of just one of her beliefs. She would, accordingly, have to also believe that "Either a cat or a dog is making the commotion in the bushes, or 2+2=7.", as salva veritate principle still obtains, even though the second disjunct is obviously false.

The problem of logical omniscience, as we have previously briefly discussed, is a simple consequence of a system's deductive closure under implication and its direct application to the domain of epistemology. This application results in *epistemic closure on material implication* and usually takes form of a *modus ponens* instance; If an agent knows that φ obtains, and she knows that if $\varphi \rightarrow \psi$ also obtains, then she ought to know that ψ obtains, mutatis mutandis for belief. However, seeing that the conception of belief appears to imply much more awareness of the proposition than knowledge does in our informal epistemologies, it seems that the account of logical omniscience for belief sounds more natural an non-problematic to some extent. At first, we might be more keen on accepting the following argument 'If I believe that φ, and I also believe that $\varphi \rightarrow \psi$, then I believe that ψ.", rather than its epistemic iteration. Nonetheless, as our doxastic logics validate the rule of necessitation, it is quite easy to see that the problem is just as serious, notwithstanding our intuitions.

I have noted that numerous authors (Fagin, Halpern, 1988; Rantala 1982; Priest, 2005; Pacuit and Salame, 2004) that work in the field of formal epistemology often resort to various attempts of weakening deductive closure of the system in order to avoid the undesired consequences of its epistemic rendition. However, as we have seen, this comes at a price of compromising some valuable properties of the systems we are taking into account. With the introduction of logics weaker than normal (meaning any logic that invalidates either the Axiom of Distribution or the rule of Necessitation), some of the principles we wanted to preserve might no longer be validated. This puts the modeller is between Scylla and Charybdis of either having to weaken the logic they use for the model, or accepting the problem of logical omniscience as a consequence of classical closure and its epistemic application.

I have previously stated in the introductory chapter of the dissertation that I will attempt to show that there is another way out of the logical omniscience. This proposed solution will be presented in the following chapter of the dissertation that deals with Distributed Systems Models and algorithmic knowledge. This way out might be contentious because it integrates the PLO into the logic for knowledge, but through the process of algorithmic dynamisation of inference reframes the problem in an uncommon way. The logic of knowledge remains normal with appropriate properties, however, it only serves as a framework for all the propositions that are in some way inferable within the system. As opposed to the framework of inferability, it proposes a more stringent dynamic notion of computability. Computability itself implies the temporal structure of system, and allows the inferences to be defined differently than a mere set of relations, which was the case in the static system we have thus far observed. The following chapter will, hence, elaborate further on the framework of DSMs and will attempt to clarify the difference between knowability and computability.

Within his research, Stalnaker (1999) has considered modelling S5-defined knowledge within the DSMs, however, it might be proven useful to see if some weaker logics are applicable to the DSMs. DSMs' usual epistemological interpretation goes a long way in setting the stage for an externalist approach to knowledge, which I intend to defend. However, as Stalnaker is well aware, his analysis goes to show that no epistemic model based on a normal epistemic logic will have an out with the PLO, but some attempts to clarify it further can be made. In the next section I intend to show that, while Stalnaker's endeavour does the job for constructing a thoroughly realist formal epistemic theory in discussing accessibility of information within the given framework, there is an alternative approach.

DISTRIBUTED SYSTEMS AND EPISTEMIC EXTERNALISM

This part of the chapter pertains to a framework for developing epistemic and doxastic logics in a dynamic, multi-agent environment. The Distributed Systems Models are used in the literature (Halpern, et al.,1994, Stalnaker, 1999) to capture various formal properties and philosophical problems of epistemic and doxastic systems. One instance of explicating philosophical and formal problems with this framework is Stalnaker's (1999) endeavour of clarifying the problem of logical omniscience through a set of epistemic situations displayed using the inventory that is defined for Distributed Systems Models' structures. The basic framework which will be defined throughout this chapter is highly idealised with respect to the type of agents that it deals with, however, it will hopefully become clear that we are able to enrich the system's structure with formal instruments that should work in favour of hindering this idealisation. As is the case with scientific and formal models in general, the more specific we get in defining the instruments of the system, the intended domain of application becomes narrower. I will attempt to show that this is not necessarily the case here, as the intended domain of application will not become narrower, but the description of the constituents and their behaviour will become more nuanced. This is the case because we will be able to observe them in a finer-grained environment, without specifying their properties and functions that will describe their programmability for action. The Distributed Systems Model itself will finally serve as the basis for applying the set of logics that should describe the theory of knowledge and belief that I decided to put forward in the dissertation.

As this constitutes the final part of the introductory chapter on logics of knowledge and belief, I will provide a brief narrative and formal explication of their inventory within a static epistemic and doxastic formal environment. I will address the components of the system, the rule-governing structures, and will give some examples which should make this introduction clearer. The dynamics of the Distributed Systems Models will be added in the next chapter when I introduce the notion of algorithmic knowledge and programmability of agents within the system for action.

DSMs were primarily introduced in the field of computer sciences and were used to describe the behaviour of various computer components that interacted within an interconnected system, i.e. a network. Their respective behaviours could are often described in terms of simultaneity within system, as one needs to be able to ascribe epistemic and possibly doxastic states to components at a certain point in time, or otherwise a specific permissible situation in the system, but this dyamisation will be introduced later on. A permissible situation in a system is basically a set of formal parameters for the constituents that are regulated by the logical-epistemic apparatus governing the system.

Stalnaker introduces the notion of DSMs as a means of epistemic modelling in order to describe and elaborate the interrelation between knowledge, belief, and action. As he sees it, as theorists of knowledge, we are not to strive to deal with the mental states of agents in order to grasp whether the said agent is in the state of knowing or believing something. Rather, we infer to agents' knowledge or belief by observing their behaviour, as Stalnaker (1999 p. 260) asserts that most of our respective knowledge and belief is manifested through action. This basically entails a sort of behaviourist and externalist enterprise which is not necessarily language-based. It does not itself exclude linguistic structures as indications of epistemic and doxastic states of agents, it just does not depend on them exclusively, as is the case with most other syntax-oriented epistemic models, such as the Sentence Storage Model (Stalnaker, 1999).

This type of modelling also has an advantage of pertaining not only to human agents, but to a wide array of agents that can range from being as simple as a light switch or a neuron to being as complex as a country's economy. For example, within a DSM, an economy can be viewed either as an agent or a system, depending on how we choose to frame it. If we were to conceptualize it as an agent, then we would view its behaviour in relation to other interconnected economies with discrete possible states, and we would be justified to ascribe it relevant properties in order to anticipate its further behaviour. We would also have a clear understanding of how its variables depend on variables of other interconnected entities in the system. Conversely, if we were to view an economy as a system through DSMs, then we would have to formulate its own constitutive elements and observe how they interact between themselves. We could, as a gross oversimplification, view the economy as a system of agents that exchange resources in a quid-pro-quo manner. The attribution of relevant features such as knowledge and belief would reflect the modal character of their respective states that are congruent with the rules of the game.

So, in order to get a grasp of what a prototype of a DSM looks like, we can turn to Stalnaker's (1999) paper "The Problem of Logical Omniscience II". To shortly sketch out the proposed framework, let us consider a system of *interconnected processors* (P_0, P_1, P_2, ..., P_n). The processors can be understood as epistemic and doxastic agents in this abstract scenario which can carry information. One of the processors we can understand as a null-processor; a simple unit displaying some state of affairs in the environment. It is incapable of being in a specific epistemic or doxastic state, as it solely works as an input feed of external information for the system. As the processors are interconnected, they have some insight as to what is going on within the system, but most of their knowledge is inferential, as I will soon elaborate further.

Each of the processors is capable of being in a finite number of mutually exclusive, but jointly exhaustive states [s_0, s_1, s_2, ..., s_n]. We can refer to such states of each individual processor as *local states*. Now, a set of all local states of a system, exactly one from each processor is called *a global state* (g_1, g_2, g_3, ..., g_n). A global state, hence, comprising all the *actual* local states within the system, can be viewed as a specific situation, or rather a scenario of the system at some point in time. As global states of the system shift, so do the local states of individual processors.

Observing the behaviour of the processors within the system, we can infer to the changes within the global states of the system. For instance, a processor P_0 can be in a local state S_1 only if the admissible global state of the system is either g_6 or g_7. This is carrying information to the observer of the system that there are only two global states in which the system is permitted to be in at the moment. Now, if we were to observe another processor, for example P_2, and we see that the processor is presently in the local state S_9, which is admissible only if the global state is either g_7 or g_{20}, then we can infer that the global state of the system at this time is in fact g_7, as it is the only element in the intersection of sets of admissible global states with respect to the local states of the processors.

In other words, we can infer to global states of the system in a similar manner in which Jane from our previous example could infer to the location of her actual world within the frame of possible worlds that are doxastically or epistemically accessible to her. This being the case, we are still discussing only the information that is accessible to the theorist. We have still not ventured into the explanation of how Jane made her decisions in terms of reacting to some external stimulus. So, it appears that we are able to talk about two distinct kinds of information availability; (1) the information that is accessible to the modeller about the epistemic and doxastic states of the processors in a distributed system, and (2) the information that is accessible (and usable) to the processors within the system when they are in a certain local state. Both kinds of availability are hindered by the problem of logical omniscience, but for different reasons.

So now the question imposes itself as to how can we know that a certain processor is in a certain local state. As we have already discussed, the respective epistemic and doxastic states (or in the terms of DSMs local states of processors) are manifested through action. Under the supposition that Jane responds to evidence adequately and follows the rules of logic non-erroneously, Jane's decision to either remain seated or get up and swiftly move from the bench is the relevant indicator for inferring to her epistemic or doxastic states. But all of this is manageable without the apparatus of DSMs. So, what exactly do we get by introducing such a way to talk about belief and knowledge attribution?

The simplest answer to this question is that DSMs help us flesh out some contentious properties of modal knowledge and belief attribution when describing *information location* (Stalnaker, 1999, p.260). DSMs allow us to clearly and analytically delineate the type of knowledge that processors possess, thus describing a global epistemic situation pertaining to the system in question. Or in simpler terms, we can assert where the relevant information is located within the system. A proper description of information location within a system can basically tell us if the processors' knowledge is mono-agent, common, distributed, first-order, second-order, etc. A good exemplification of this is the muddy children puzzle.

As a brief sketch of muddy children puzzle can be construed as following: Suppose there is a number *n* of children playing together in front of their house. It was raining and their mother told them not to get dirty, otherwise they will face punishment. The children are all very shrewd and prudential, but they'd all enjoy seeing their siblings getting caught for misbehaving. Now let us assume that a number of them, *m*, got mud on their foreheads. They obviously cannot see their own forehead, but they can see others'. Now their father arrives, and proclaims "At least one of you is muddy on your forehead." He then asks them repeatedly, while allowing them to answer and for them to hear each others answers :"Are you the one with mud on your forehead?". It is easy to see if *m* is two or greater, no child will respond affirmatively. But here is the catch – prima facie, the father didn't provide any new information to any of the children. All of them already know that at least one of them has a muddy forehead. But only on the k^{th} the father asks "Are you the one with mud on your forehead?", the answer of the child with mud on their forehead will change. This is the case because even though the children shared the knowledge that at least one of them had mud on their forehead – common knowledge, they did not share this kind of distributed second-order knowledge about what their siblings knew based on what they saw and what they were able to infer from each others answers. It appears that the father's proclamation, which each of the children heard, and furthermore, each child's knowledge that each of them heard the proclamation, are they able to make an inference about them being the one with mud on their forehead. These different kinds of knowledge are usually delineated by the terminology mono-modal knowledge – pertaining to one agent, and multi-modal – pertaining to multiple agents within a shared situation.

For now, return to our simpler example with Jane and the dog in order to show how we can frame it within a DSM. We can know that processor JANE's decision on whether to move or remain seated is based on the second-order epistemic or doxastic state of the information that a cat or a dog jumped out of the bush – the null processor's local state corresponding to a fact in the environment. Our null processor, let us call it BUSH is capable of being in two discrete mutually exclusive and jointly exhaustive local states; 'dog' state and 'cat' state. As JANE is the only active processor in the system, we can see that her two local states, programmed into actions as 'stay' and 'leave', correspond bijectively to the local states of BUSH processor 'cat' and 'dog'.

Hence, we can infer to JANE's local states from observing BUSH's local states and vice versa. Furthermore, not all global states are admissible in this DSM. As we have seen, the bijective function between local states of JANE and BUSH essentially prohibits them from being in a global state, let us call it G3, where JANE is in 'stay' and BUSH is in 'cat'. The same is obviously the case for the global state G4 where JANE is in 'leave' and BUSH is in 'DOG'. The other two global states are admissible, as can be understood from the biconditional form of the scenario.

CHAPTER IV – DISTRIBUTED SYSTEMS AND ALGORITHMIC KNOWLEDGE

KNOWLEDGE AND ACTION

In the field of formal epistemology, generally speaking, attibution of knowledge can be two-fold. In an externalist formal epistemic model, one layer of understanding and explicating knowledge ascription is based on information that the modeller has at disposal by observing the processors (system components, agents) interact. In other words, the knowledge that he has of their abilities to (successfully) communicate or interact with their environment is solely based on a behaviourist theoretical framework, as the ascription of epistemic and doxastic states to them is grounded on their local states corresponding to some state of affairs in the world. More precisely, it is grounded on observing actions indicating their being in a particular local state capable for determining such actions. As a sidenote that I will delve into further in the chapter on verificationist modelling, this itself does not necessarily corroborate a realist epistemology, as 'the world' can be framed in the way to mean 'all demonstrable truths'.

The idea of using intensional semantics for dealing with knowledge ascription in an externalist model is appears to be very natural in the context of dealing with complex multi-agent systems where the notion of knowledge is somewhat technically revised and can encompass theoretical variants such as common and distributed knoweldge. This kind of framework was often used in the fields of game theory and decision theory, as the formal explication of the concepts such as rational choice or higher-order knowledge have indeed been shown to be most valuable in these kinds of analyses.

As I have already discussed in the chapter about adding dynamics to epistemic models, DSMs have been the golden standard of approaching multi-agent knowledge and belief ascription based on action. As a reminder, it is important to note that DSMs do not necessarily rely on language, as most of what agents know or believe is demonstrated through their actions and interactions. But herein lies the gap in the explanation - the epistemic and doxastic asciptions do not at all rely on what the processors are permitted to infer with regards to the system in use in order to problem-solve or make decisions. For that gap to close, we ought to take a closer look at what procedures are, in fact, available to the agents that are making the relevant choices.

The notion of an algorithm was briefly introduced when dealing at the begnning of tthe chapter, but it appears that we are in need of a specific formal explication of what precisely constitutes this ability of computing optimal choice-making behaviours with respect to what is known or believed by agents within the system. To the proposed elaboration of the constituents of DSMs we can add some notions that should account for the models' dynamisation and establish the groundwork for their behaviours through time.

As a basis for this development, I use the paper Algorithmic Knowledge written by Halpern, Moses, and Vardi (1994). I mainly adopt their proposed terminology with some slight modifications in order to work out a consistent and unambiguous nomenclature with the remainder of this disertation.

So far, we have defined a set of processors $p_0 \ldots p_n$ (one or more, which are NULL processors, working as an input channel for the system, and hence displaying some state of affairs in the environment), a set of local states of each processor $s_0 \ldots s_n$, and a set of global states of the system $g_0 \ldots g_n$. These elements were sufficient to display some possible states of the observed system, and to establish a set of formal relations between the local states of the processors, and their knowlegde with respect to some input from the NULL processor.

DYNAMISATION

Now, in order to lay out the groundwork for adding dynamics to the system S, i.e. to show how it would behave given a temporal dimension, we add the notion of a run, $r \in S$. A run is conceptualised as a sequence of global states, $r(t)=g_0, g_1, \ldots, g_n$, in one possible execution of the system. This sequence of global states can be read as a possible sequence of events (or even more precisely – situations) that occur in some context. If we would situate this idea into the example that I used in the chapter on epistemic and doxastic systems, we could say that the first global state within a run is Jane arriving at the park with her dog, when the bench is still empty and the bush being silent. The second global state could be interpreted as Jane sitting on the bench with the dog by her side, while the bush is still silent. The third could be understood as bush starting to shake with Jane still sitting on the bench, and so on. We interpret the relevant objects in the system as processors, JANE, BUSH, BENCH, etc. As we have seen, each of the processors has the capacity of being in multiple mutually exclusive and jointly exhaustive states (BENCH can be empty or occupied, BUSH can be silent or shaking, JANE can be sitting or leaving). In this case, JANE is the only active processor, meaning that she is the only part of the system capable of having epistemic and doxastic states, while the rest are NULL processors, i.e. the environment. The run is then viewed as a sequence of these states, or possible situations that can follow one another. The impremissible states are then situations that cannot follow one another as such sequences would not abide by the rules of the system or simply aren't intensionally accessible one from another. For instance, it would make little sense if Jane was first sitting on a bench in a park, and then she arrived at the park. The latter is necessary condition for the former to make sense. So, the run is then a set of scenarios that can play out, given the starting conditions of the system. As JANE is the only programmable (even self-programmable) processor, her actions can be viewed as being responsive to evidence, i.e. she will make decisions based on the behaviour of the system as a whole, while the theorist observing the system will be able to ascribe epistemic and doxastic states to her based on examining her actions.

A set of all permissible runs r in this framework is precisely how the system S is defined. Furthermore, a run r is paired with a temporal point t that gives us the information about the present global state of the system. So, it is obvious that in order to determine which global state a system S is in at some point in time, we ought to have access to both pieces of information (r, t), as the run determines the sequence of global states, and t defines the point in time at which the system is observed.

A local state of a processor is then defined as $r_i(t)=s_i$, as the local state of a processor depends on the global state of the system S, which is determined by the pair (r, t). As the processors have no way of telling which possible run of the system is currently in execution, nor do they possess the knowledge of what point in time they are in in terms of one of the possible scenarios playing out, we can easily see that two states of the system (r, t) and (r', t') are indistinguishable to them if they are in the exact same local state. This relation of subjective indistinguishability is usually denoted with the symbol ~. This indistinguishability relation itself does not give us any information about the logic that the model validates, even though the Distributed Systems Models usually validate an S5 logic. The said relation behaves as an S5, universal accessibility relation, regardless of which logic of knowledge and belief we assume for the system to use.

We say that $(r, t) \sim_i (r', t')$ iff $r_i(t) = r'_i(t')$, as the processor's local states are equivalent at both points within the system's execution.

As the dynamic system can now be syntactically and functionally complete, the only thing remaining is to define the semantics for it, which is easily achievable by introducing a simple valuation function $[\![\ ]\!]$ (Halpern et al., 1994) decided on naming it the mapping function π), which basically assigns a truth value to any known proposition φ at a certain point in time within a given run. With the semantic valuation function added to the system S - $(S , [\![\ ]\!])$, we now have an interpreted system I. Within the interpreted system I, we can say that a processor knows some formula φ (Kiφ) in the run r at the time t if φ is the case for every (r', t') that are indiscernible from (r, t) to the processor. Formally, we can display this as following:

DEFINITION. $(I, r, t) \vDash K_i\varphi$ iff $(I, r', t') \vDash \varphi, \forall(r', t') \sim_i (r, t)$.

This will represent our notion of *implicit knowledge*, i.e. the knowledge that our processors possess within the framework of the logic that the system validates. This notion of knowledge corresponds to what Stalnaker (1999) refers to when he states that the model is *externalist*, as this sort of knowledge ascription is establishable without the theorist needing insight to the computational capacities or storage space of the participants of the system. As I have already stated, it is the knowledge that is *in principle* inferable by means of the logical apparatus that the model validates. However, as it should be quite clear by this point in the discussion, not all knowledge is computable, even in optimal circumstances. So, in order to delineate between this idea of possible, implicit knowledge, supported by the logic that is in use, and de facto computable knowledge, Stalnaker (1999) and Halpern et al. (1994) here introduce the notion of *algorithmic knowledge*. Although they propose a slightly different terminology (Stalnaker usually refers to a 'selected feature of a local state), they generally consider the following idea: for some formula to be considered as known, or even knowable for the processor *i*, there has to be an explicitly describable procedure for telling whether the formula is true or not.

Halpern et al. (1994, p. 285) rightly note that having an algorithm for computing an output is not sufficient for establishing that an agent has knowledge of a proposition φ based on the output alone; they state that knowing *when* to use it is just as important a part of the equation. They propose a situation in which a processor has two trivial algorithms, one of which answers "Yes", while the other answers "No" to any question indiscriminately. Now, if the logic that the model validates is decidable, then each proposition within the system is assigned either truth or falsity. This would mean that such a processor has correct answers to any question it is asked, as it possesses at least one algorithm that generates the correct answer. However, it is not stated that the processor has at disposal the knowledge of when to implement the "Yes" algorithm, nor the "No" algorithm. Thus, as the authors put it; "Part of 'having' an algorithm is knowing when to use it'. This is precisely the reason behind theoretically incorporating algorithms within the local states of processors, as they are now bound to specific runs and moments in time of the execution of the system, by which we can ascertain if the algorithm that was implemented in generating an answer to the question actually represents the *disposition* of a processor to be susceptible to evidence that surrounds it. In other words, the processor has to be sensitive to its environment in a particular way that allows it to interact with its environment adequately.

From a formal standpoint, we define the local state of a processor that possesses an algorithm as an ordered pair (Alg_i, $data_i$ (r, t). For brevity, we can refer to Alg_i as A, and $data_i$ as c for context. We can define A (r, t) as the algorithm within the run r, at a moment in one of system's possible executions t, and a context – c (r, t), that is basically the remainer of the local state. Context can, hence, comprise the formally defined $data_i$(r, t) as all of information that the local state possesses. The authors denote this formal setup as an interpreted algorithmic system. It might be noted that there can exist vacuous local states of the agents within the system that play the role of states that the agents are in when they are not responsive to some intrasystematic variables of their environment, i.e. not being responsive to some evidence, i.e. input.

Now, in order to formally capture the notion of knowledge that is *actually* accessible to the processor at a certain point in time during one of the possible executions of the system, we ought to introduce a new operator for it. So, algorithmic knowledge, i.e. computable knowledge of the agents will henceforth be denoted with the epistemic operator $X_i.\varphi$ This operator can be read as ‘the agent i algorithmically knows φ’, or ‘the agent i explicitly knows φ’. It serves as a ‘stopper’ for logical omniscience, as it is now expressible within the system that the agent has an algorithm for computing φ, while φ implies ψ, but it need not to possess an algorithm for computing ψ, even though ψ is in principle inferable from these two facts with the logical instruments at disposal, i.e. even though ψ an intrasystematically inferable in the context of *implicit* knowledge.

So the authors (Halpern et al. 1994, p. 259) propose the following explication of algorithmic knowledge operator for this model:

DEFINITION.

$$(I, r, t) \vDash X_i\varphi \text{ iff } A\langle alg_i, c\rangle\ (r, t) = \text{“Yes”, for } alg_i\ (r, t), \text{ and } c = data_i\ (r, t)$$

The notion of local data incorporated in this model can be interpreted as an awareness function from any $K_i\varphi$ (implicitly known φ) to some moment in time within a possible execution of the system (r, t). In other words, any information that the agent is aware of at a certain point in time in a run is indeed available information for potential inferences when applying an algorithm. Everything up until now sounds great in theory, however, a looming problem that Halpern et al. (1994) also bring to readers’ attention is that none of this, in fact, *guarantees* that the computed information that the processor provided as output is even correct. Algorithms that we use in our everyday lives are not infallible – we infer wrongly when presented with evidence, we make inaccurate assessments of how much time we need for a task, we even introduce false beliefs into our belief sets for various extralogical reasons such as emotional states and similar motivating factors. However, there is a way to formally describe algorithms that do their job adequately, i.e. algorithms such that $X_i\varphi$ implies $K_i\varphi$. Sound algorithms can be explicated in the following way; Any local algorithm that an agent i possesses at any point in time during a possible execution of the system (r, t) is understood as sound for the interpreted system I and formulae φ if the local state with the embedded algorithm provides an output “Yes” only when $(I, r, t) \vDash K_i\varphi$, and *mutatis mutandis* for the “No” output. The authors note that if such conditions are met, then the newly introduced $X_i\varphi$ is closed under factivity, i.e. the Axiom T, stating $I \vDash X_i\varphi \Rightarrow \varphi$.

Since we are trying to establish a basis for the application of DSMs to our ordinary knowers, here we are attempting to reconcile the idea of capturing these highly idealised universal procedures for ascertaining the truth of some proposition with some examples of ordinary knowledge. Tennis player who aptly runs down every ball during a point returns it with great precision and power, while not having the ability to explicate or describe specifically how they did it, or the detective who believes somebody is lying while testifying without necessarily knowing why they believe it. The tennis players most certainly cannot supply the mathematical description of the point in question, including the variables such as the spin of the ball in rounds per minute, the speed of the ball in meters per second, and the angle of the ball at some moment in time. The detectives might have no general procedure for telling if someone is genuine, but can still successfully get to the truth of the matter.

As we have seen, that the defined accessibility relation for the system, $\sim_i$, is an equivalence relation that is specific for S5 systems. As we have briefly already mentioned before, DSMs were originally developed in the field of computer sciences, which means that the model of knowledge they were inclined to adopt didn't have to closely mirror or even pertain to our ordinary conception of knowledge. It was in fact quite far removed from it, however that posed no issue whatsoever, since the DSMs were never originaly intended for the usage of modelling knowledge in our ordinary sense of understanding it.

However, with the introduction of the algorithmic knowledge operator and its logical infrastructure (both syntactic and semantic), we have pushed this model closer to our common unidealised inference-makers. It is also of importance to remind the reader that the entirety of the proposed apparatus is *not* in tension with using a logic for knowledge weaker than S5. In this dissertation I hope to have shown that the S4.2 system could be proven to be a better candidate for the type of modelling this dissertation is focused on – the common knower. Any of the principles used for explicating the notion of algorithmic knowledge are still covered by the structure used in S4.2 logic of knowledge, with an exception of subjective indistinguishability which remains closed under the Ax5 of negative introspection. Subjective indistinguishability, as discussed in the first two chapters of the dissertation is the superset containing both sets that are defined via the accessibility relations for knowledge and belief. Again, this is the case because out of all subjectively indistinguishable sets of worlds, only some are in the set of worlds consistent with what is known (including the actual world if we accept factivity of both knowledge and algorithmic knowledge), and some are in the set of worlds that are compatible with what is believed. As a reminder, if the sets of worlds compatible with what is known and those compatible with what is believed coincide, then we have no false beliefs. Otherwise, if they do not overlap at all, then all the beliefs that an agent has are false. But there might exist worlds that are incompatible with both what is known and believed that appear to be indistinguishable to us. Even though our present formal structure of knowledge and belief does not conform to this universal accessibility relation, we can still explicate it as a useful tool for understanding the hierarchy of the model in question.

KNOWLEDGE ACCESSIBILITY

We are now certainly aware that normal epistemic logics suffer from the problem of logical omniscience. When we consider DSMs as a possible application of such logics, it is clear that the processors within the distributed network are omniscient. This is of no concern to the game theorists' analysis, as they have little problem adopting an S5 system, with its contentious axiom of negative introspection. The axiom states that if an agent does not know some proposition φ, then it knows that it doesn't know it. In terms of the formalism presented in the field of game theory, this is quite common, as we are not interested in any „real" knowledge or rationality of agents, but only in the optimal computational outcome as a solution to the given problem. So far, we have seen in both Halpern et al. (1994) and Stalnaker (1999) that in order to even scratch the surface of the notion of real knowledge, even in such a highly idealised environment, we must be able to differentiate between implicit knowledge – the one that the logic that we use validates, and computable knowledge that we acquire by implementing an adequate algorithm for computing the output of the system in question. As a reminder, canonically, computable knowledge is often refered to as algorithmic or explicit knowledge, with Stalnaker even using terminology of avaliable or accessible knowledge. Generally speaking, it not only gives us insight into what is (or can be) known, but also how it is known.

The notion of available or accessible knowledge for multi-agent contexts such as Distributed Systems Models is quite useful when trying to resolve the Problem of Logical Omniscience. As a part of his analysis, Stalnaker (1999) proposes several theoretical maneuvres that one can take in order to formulate a non-trivial and precise conception of computable knowledge for the participants of the system. Admittedly, his work on the PLO reached no ultimate solution to the problem (at least for the normal epistemic systems), but has adequately clarified the groundwork for potential approaches.

Firstly, it might be useful to reconstruct a variation of one of the examples (Stalnaker, 1999) he used that I believe encapsulates this problem well and fleshes out the issues that arise when attempting to provide a clear distinction between implicit and algorithmic knowledge. Let us suppose that we have a 16-bit processor, which means it has around 65000 possible local states. In principle, each of the states is programmable for action by means of possessing an internal feature – an algorithm – which is capable of determining some processor's action on the condition that some proposition φ obtains. In other words, the processor should, if possible, perform an action if φ is the case, and perform a different action is φ is the case. Performing such actions means that the processor is in some way calibrated to compute the truth of φ at a given point.

The said action can be a basic output function that maps out the answers 'Yes' and 'No' to the input variable. Now, let us suppose that the input in this example are integers. Our 16-bit processor has to be disposed, i.e. possess an algorithm, that outputs 'yes' whenever (and only when) the given input integer is prime, and output 'No' whenever (and only when) the input integer is a non-prime. Let us assume that each interger that is given as an input to the processor corresponds to a single local state that the processor has, say, from 1 up to around 65000 thousand in the case of our 16-bit processor. Now, for instance, given the input 1447, the processor will now have the implicit knowledge that the input integer is prime, as this property of the number 1447 is bound to the internal property of the local state 1447 of the processor which is active only when the input number is composite. However, Stalnaker notes that no action might be programmable for this input, as the information about the number being prime might exist, but it might not be accessible to it as it might lack an operational algorithm to compute it. We might even consider the case in which such an algorithm for computing if φ obtains actually exists, but it might generate a too complex computational task for our processor.

Thus, we might want to consider only the features of local states that are actually usable for our processor – it will only have the knowledge that some proposition φ obtains at a certain point when and only when (1) information is, in fact, there, and (2) there is a simple enough way to access it. The criterion that we propose for an algorithm to be usable for the processor, however, has no effect on what information is there to begin with. In other words, the epistemic logic that we use will determine what information exists within the system, the model will determine its location within the system, and the criterion for selection of the algorithm will determine in which way the local states are programmable for the processor to take action under the condition of some proposition obtaining.

CHAPTER V – BELIEF REVISION AND EPISTEMIC UPDATE

The systems that we have described so far, S4.2 for knowledge and KD45 for belief have been shown to possess some desirable properties that we have been discussing so far such as completeness, soundness, consistency, and decidability. However, in virtue of modelling some theory of knowledge and belief applicable to the intended domain of our ordinary knowers, one crucial element was left out – our agents have to be able to be wrong when believing something, and they must have a way to manage it in some well-structured and clearly definable way. As Negri and Pavlović (2023) rightfully notice, the vast majority of the theorists agree that there are at least two non-contentious elements in discussing knowledge. The first is that it entails truth – we cannot know something false. The second element is that knowledge implies belief – we cannot know something we do not believe. However, it appears that there exist many situations in which we have true beliefs, but wouldn't be keen on attributing ourselves knowledge, e.g. we can guess correctly or have an unfounded or ill-founded true belief. So there is an evident need for expanding the conception of knowledge as true belief in some non-trivial way in order to arrive at a definition that would possibly come closer to our folk understanding of what knowledge is. Negri and Pavlović (2023) finally settle on Sosa's Epistemic Defeasibility augmentation element after shortly commenting on Gettier's and Russell's counterexamples to the well-established definition of knowledge as a justifies true belief.

The following chapter will, hence, focus on what is called 'The Dynamic Turn' in the field of formal epistemology and it should provide us with the necessary tools for accounting for revision of belief states and updating our knowledge database. The idea of adding dynamic elements to our static models stems from the original basic belief revision theory AGM (Alchurron, et al. 1985), and most contemporary variations of belief revision and epistemic update use it as a framework for developing specific formal mechanisms which were not included in the original iteration. As was described succinctly in van Benthem (2007, p. 1), firstly, the original AGM theory and later AGM-style theories are almost indiscriminately mono-agent accounts dealing with factual information only, while providing only an abstract framework for what would ideally be a syntactically fully developed theory. As we are interested in multi-agent systems in which the said agents are capable of interacting, having higher-order beliefs about their own states and the states of other agents, we appear to be in need of some more structure.

In this chapter, the plan is to sketch out AGM and its postulates, after which I will attempt to delve a bit further into logics of belief revision and epistemic update that offer a specific set of procedures that should describe the behaviour of agents' epistemic and doxastic states when engaging with new external information. Van Benthem (2007) himself provides an embedding for the AGM theory with the system of Dynamic Epistemic Logic

(DEL for short henceforth) which specifically explicates procedures of epistemic update through the mechanisms of public announcement. After I analyse the pros and cons of DEL embedding of AGM systems, I will revisit another attempt for dynamising the static logics of belief and knowledge we have in order to avoid some of the unwanted consequences of adopting DEL-style systems. On the opposite side of the dynamic spectrum of doxastic logic, we can find theories such as CDL (Negri and Pavlović, 2023). Conditional Doxastic Logic is not in itself dynamic, and is partially supported by a Kripke frame (meaning that one of its relation of accessibility is supported by an S5 frame, while the other is supported by a neighbourhood fame), however, it is capable of describing the relevant dynamics through a set of relations between the sets of what is given (a description of a doxastic state) and what is accepted on that basis (if anything should be added or changed within the set of doxastic states). At the end of the chapter I intend to discuss why CDL has proven to be the most promising choice of the model I have been developing as the main thesis of this dissertation, but also why we can always easily calibrate the model for PAL and DEL systems and their accessibility relations if the need arises. It can also be shown that such calibration can be done without compromising the model's soundness, completeness, and decidability.

AGM-STYLE SYSTEMS, PAL, DEL, AND THEIR PROPERTIES

First let us start a brief overview of (1) AGM theory of belief revision, then we will briefly delve into Public announcement logics (PAL), and finally, (3) DEL embeddings of the AGM system.

AGM-style theories operate under a simple hypothesis that states that our beliefs are not always correct, complete[4], or consistent. Thus, the theory offers us a frame for developing an apparatus for righting those wrongs. The structure of the theory is divided into three operations, viz. (a) expansion $\oplus$, (b) contraction $\ominus$, and finally (c) revision $\circledast$. The basics of using the three operations are closed under the three sets of postulates, one for each, which establish the behaviour of the set once some element has been added, removed, or revised. As we are not discussing AGM in depth, but only as an introduction to the notion of belief revision, I leave the reader to possibly examine them further in the book by van Ditmarsch et al. (2008) entitled Dynamic Epistemic Logic. For now, I will only address the formal operations and their respective consequences of application.

So, we read $K \oplus \varphi$ as the set of beliefs expanded by φ. With regards to the postulates of expansion operation, that means that the proposition φ has been added to the belief set

[4] As a point of clarification, 'complete' here does not refer to the relatioship between the beliefs' syntax and semantics, but only to the cardinality of the belief set in the context of offering a full description of the observed phenomenon.

with minimal revision. In other words, when we introduce a proposition into our belief set, we will add as little new assumptions as possible with respect to the existing elements. Furthermore, we read K ⊖ φ as the set of beliefs that was contracted by excluding the formula φ from the set. The problem addressed by van Ditmarsch with respect to contraction is that there might exist a formula in the belief set that depends on the excluded formula φ being true. Under closure of consequence, that would imply that by relinquishing one belief, we might be relinquishing other as well without explicitly allowing so. Here the authors introduce the concept of entrenchment, which allows the agent to prioritise some beliefs over others. The final introduced operation is the one of revision, K ⊛ φ, by far the most ubiquitous in relevant literature. It operates under the assumption of calibrating the set of beliefs to newly acquired information, and hence serves as the basis for the newest developments in dynamic formal epistemology. Both DEL and CDL primarily deal with this notion of explicating the dynamics in the behaviour of the system once something new has been introduced or something old revised. So, we move on from the operational basis of AGM-style theories into the two contemporary iterations of the dynamic shift in epistemic and doxastic logics. As I have announced, we are now moving toward DEL logics which naturally incorporate PAL operations into their syntax while embedding the AGM systems postulates.

Public announcement logics are the basis for developing dynamic systems that are event-based, which basically means that they have an implied temporal dimension, or at least a sequence of states or events that provide the theory with the tools for grasping change in the doxastic and epistemic states of the participating agents. A well-known example that is often used for elaborating PAL-based structures is the muddy children problem that I have discussed in the second chapter on epistemic and doxastic systems within the introduction to Distributed Systems Models. Public announcement systems use a newly defined dynamic operator, usually marked with a dynamic variable !P, that stands for a proclamative event from a trusted source[5] that took place in some epistemic or doxastic situation, pertaining to the agents of the system. Note that !P itself does not denote a proposition that was introduced as the input for the system but the *event* of stating some new information to the participants of the system. As seen in Benthem's (2007) paper Belief revision and dynamic logic, PAL systems are often capable of defining several useful formal instruments capable of capturing mono-agent knowledge, common (or group) knowledge, and action expressions. The event of public announcement by the null processor is defined by action expression that establishes the operative dynamic element of the system. Here I propose a brief explication of the language of PAL systems in

[5] In DSMs vocabulary, 'a trusted source' can be viewed as a NULL processor that feeds true information into the system of active agents. The example that we observed in the second chapter used 'dad' as the NULL processor, tacitly implying that he always tells the truth to the children attempting to determine some information about themselves and their environment. In many realist readings of DSM modelling (Stalnaker 2006, Fagin, et al. 1995, etc.) the input feed is presumed to be correct because it *is* the state of affairs in the world.

Backus-Naur form based on Benthem's (paper) with minor appropriate revisions, in order to maintain cohesive terminology within the dissertation.

DEFINITION.

$$\mathscr{L}_{DEL} ::= p \mid A \mid \neg A \mid A \wedge A \mid A \vee A \mid A \supset A \mid K_s A \mid C_s A \mid [AEX]\, A$$

$$[AEX]: !P$$

The reader can notice that the language introduced is very similar to what we have already observed in the previously addressed systems in this dissertation with the exceptions of the common knowledge operator C_sA pertaining to a formula that is shared between all the participants within the group, and by action expression [AEX] denoting the public announcement of a fact to the participants. Such vocabulary was primarily introduced to be optimised to deal with multi-agent structures, as the contemporary standard often deals with knowledge and belief in a collective environment. This can be employed to discuss many epistemic situations in contemporary epistemology such as game theoretic applications pertain to public deliberation and decision making, democratic voting processes, etc.

Following van Benthem (2007) one can now understand the act of public announcement as possible world collapse or elimination. Let us suppose that the participants of our system had at disposal two mutually comparable epistemic cells; one in which φ was true, and another in which it was not. This dilemma was accessible to each participant at a particular moment in time. The epistemic cells were otherwise subjectively indistinguishable to the agents. As soon as the NULL processor provided the input that, for instance, φ was the case, one of the two epistemic cells collapses, i.e. is eliminated. In other words, the epistemic cell, or more specifically a singular world was eliminated by the introduction of the new information.

Formally, we can describe this for a model M, a possible world w and any true proposition φ at the world w, (M | φ, s) is the submodel of M whose domain is the set {t ∈ M | M, t ⊨ φ. Semantics of the dynamic action modality such as PAL can be defined as follows:

DEFINITION.

$$M, w \vDash [!P]\varphi \text{ if and only if } (\text{ if } M, w \vDash P, \text{ then } M \mid P, w \vDash \varphi)$$

(van Benthem, 2007).

Now I hope that it is clear to the reader that DEL system with PAL function is in its true sense dynamic. The dynamic operator captures the history of the model and displays remaining and revisited states through world elimination. This, however, is not the case for CDL system. CDL system canonically belongs to the ‘dynamic turn’ that occurred in the modern history of formal epistemology, however, it is not dynamic in the ‘true’ sense, as is the case with DEL system. This is because CDL does not have a devised formal apparatus to talk about the system’s history and world elimination through action expression. It merely *simulates* the dynamic behaviour of the system by introducing a syntactic structure that explicates the relationship between the corpus of information that is given to the agent and the belief that is its result. Hence, when describing CDL we have no need to reach for the actionable dynamic operator, as it is integrated in the structure of the system. This basically means that CDL, at least in terms axiomatically defining its accessibility relations, is essentially normal, even though one of its relations is supported by an augmented neighbourhood structure. Now, before briefly introducing neighbourhood structures that will be useful when defining accessibility of CDL, one important distinction in the epistemic theory ought to be introduced – one that regards the type of information that we are using when applying dynamic world elimination.

The majority of such logics, including DEL and CDL operate under the assumption of existence of two types of information; hard and soft (Van Benthem, 2007). The former relates to factive information that is stable and not susceptible to revision, but only positive update. The instances of hard information introduction are quite usual when dealing with factive PAL systems as they represent information that is true, i.e. corresponds to a variable from the environment. Essentially, hard information is any conclusively establishable state of affairs within the margins of the system. This notion is not formal in any reasonable sense of the word, and is most commonly used in informal epistemology. The notion of hard information directly relates to world elimination, as any conclusively establishable fact eliminates worlds in which the fact does not obtain. The notion of hard information, although primarily devised in realist epistemologies, can be adapted to an antirealist position to state that it corresponds to a sound and common source for the participating agents. The idea is well-presented in this simple example;

“The cards have been dealt. I know that there are 52 of them, and I know their colors. But I have only ephemeral beliefs about who holds which card, or about how the other agents will play. Of course, I could even be wrong about the cards (perhaps someone replaced the King of Hearts by Bill Clinton’s visiting card), but this worry seems morbid, and not very useful in understanding normal information flow. Corresponding to this distinction, different events can trigger changes in my models. An incoming public announcement !P of a fact P is a case of hard information, which changes what I know. If I see that the Ace of Spades is played on the table, I come to know that no one holds it any more.” (Van Benthem, 2007, p. 2)

Conversely, soft information constitutes a circumstantial indication that an agent has acquired while interacting with their environment. It does not constitute a fact, but rather a part of an explanation of the behaviour of the system. Taking for instance van Benthem example with cards, a smirk from a player or the look of disbelief they have once they glanced at their cards may give us a piece of circumstantial evidence that they are pleased or displeased with what they have drawn from the deck.

Now, before getting into the specifics of DEL and CDL systems, along with the framing of the latter in the context of neighbourhood semantics which are adapted to its conceptual framework, we might want to situate the distinction of hard and soft information within the context of Distributed Systems Models. The DSMs' structure and terminology allows us to get a clear grasp on what the difference is for the well-behaved idealised processors who act uniformly.[6] Firstly, another stipulation ought to be made in order for this application to work as intended; (a) the NULL processor might be an unreliable source of information, (b) it can provide the processors with partial information, or (c) the processors' competence of tracking of variables that it displays ought to be imperfect in some manner. Each of the disjuncts would imply the possibility of inconclusive evidence for a processor to act on some fact in its environment. We might be keen on adopting the second disjunct for theoretical reasons, as it would entail the problem of codifying partial information. As our system is calibrated to a propositional modal calculus, there would need to exist a way to express it in a syntactically consistent manner, which appears to be lacking when using the system that we are observing. Still, the other two disjuncts, the first and the third are modellable and can be theoretically captured within such frameworks.

Within this interpretation of the system, the hard information would constitute the cases in which the NULL processor provides a reliable and conclusive output for the processors, while the soft information would allow the processors to have indication of some state of affairs within the system without having *sufficient* information to state that it *knows* some information or its location within the system.

Thus, the remaining disjuncts that are deemed adaptable to the system can be modelled with two distinct formal instruments; either we need a (1) Bayesian fuzzy system with values determined by the interval of real numbers between 0 and 1 that will attribute the probability of φ obtaining at some point in time, or (2) we need a binary doxastic modal system with a total preorder, which will assign an ordering structure to the classes of worlds based on their respective level of plausibility. The latter is done with plausibility frames, which are shown to be interpretable in the context of neighbourhood frames. I

[6] By 'acting uniformly' I do not mean that their actions are based on the same background knowledge and beliefs, nor that they possess the same algorithms for computing knowledge. The phrase 'acting uniformly' solely pertains to the stipulation that there are no specificities from one to another in terms of their formal constitution. In other words, we are dealing with a condittional statement: "*if* they possessed the same sets of propositions that they know and believe, along with the same programming (algorithmic structures), they would have behaved the same."

will later address the interpretation of neighbourhood frames into Kripke frames with added syntactic restrictions of validating distribution and necessitation, so that the model in question is completely supported by a normal frame. So, as we are observing a formal theory that is set out to maintain system's normality, I will opt for the second option, which will then be explicated within the system of Conditional Doxastic Logic. The soft information in this reading would not constitute changing the worlds that are doxastically or epistemically accessible to the processor, but might affect the *ordering* of them within the model. This means that the processors (agents) might deem some situations more plausible to another based on the new information available in the system.

As I have stated, each plausibility frame is easily expressible within a neighbourhood frame. Not only that, but the reading of full preordering structures has a very natural reading within the neighbourhood frames, so I believe that a short introduction to them might be of use. This will allow us to capture the essential structure of Grove spheres, a type of plausibility models that categorises worlds into sets of ordered plausibility. They are usually depicted as following; The most plausible worlds, the ones that are based on hard information are subjectively indistinguishable and hence, of equal plausibility. Each superset of that set constitutes a set of worlds of lesser plausibility than the original set. I will explicate this in more detail as we delve further into the formal explication of the model and its visual representation. This might be best understood through a visual schematic interpretation that I will include in the remainer of the chapter.

NEIGHBOURHOOD SEMANTICS

I introduce the Neighbourhood frames as a generalisation over the Kripke frames for the purpose of modelling plausibility in the doxastic logic of the model. Conversely to the Kripke frames, a neighbourhood frame is not built on a linear binary accessibility relation, but on the notion of membership in a 'neighbourhood'[7]. A neighbourhood function attributes to each world a powerset of a powerset of W, the set of all possible worlds within a model. In other words, it provides each world with a collection of all of the possible subsets of possible worlds.

DEFINITION. N: $W \rightarrow \wp(\wp(W))$

[7] The term neighbourhood is often interpreted in two distinct ways – it can mean either the element of the larger subset attributed to a single possible world, or it can mean the entire subset. The distinction is often left for interpretation within the context.

Simply put, if we were to view a proposition as a set of possible worlds $\wp(W)$ (similar to what Chellas (1980) proposes, then the neighbourhood function attributes to each possible world a set of propositions that are necessary within it. The criterion of which propositions are assigned to each world is actually how we interpret the modal terms within the a neighbourhood frame. So, in terms of adding semantics to the syntactic infrastructure of the formalism, the Kripke structures interpreted modalities with the accessibility R-relations, which means that the valuation function ⟦ ⟧ assigned truth by linearly designating it to the worlds which were accessible from the world in which a formula was closed under a modal operator. So, if $\Box\varphi$ was true in the world w_0, and w1 was accessible from w_0, then we would be justified to infer φ in the world w_1. The accessibility of worlds, in this case described between w_0 and w_1 was the justification of attributing truth to φ in w_1 with $\Box\varphi$ being the case in w_0.

Conversely, the neighbourhood structures interpret the modal terms with the instrument of a truth set. A truth set is basically a set of worlds in which some proposition φ obtains. It is the operative part of interpreting modal terms within the model, as the neighbourhoods of a certain world w_0 are defined by the means of defining 'the forcing relation'. Forcing a formula in the neighbourhood model is basically what validating a formula in a Kripke model is. To say that a model in a world w_0 forces some formula A means that the neighbourhood of w_0 constitutes the truth set of A. In other words, the formula A is necessary in the world w_0, as the truth set of A *is* what defines the neighbourhood of the starting world in the first place.

Eo ipso, we define a neighbourhood model with appropriate semantics as following:

DEFINITION. $M = \{W, (N_s)_{s\in S}, [\![\ ]\!]\}$

So as we have seen, the neighbourhood function, designated with the label N, plays the role of the linear R-accessibility relation of Kripke frames. Now, as seen in Negri and Pavlović (2023) we can define a multi-agent neighbourhood, in which W is the set of possible worlds, the valuation function ⟦ ⟧ attributes each atomic proposition a set of possible worlds, for each agent s that is an element of S (the set of agents that the system defines) that satisfies the properties of

a) Non-emptiness: $\forall\alpha \in Ns(w), \alpha\neq 0$
b) Nesting: $\forall\alpha, \beta \in Ns\ (w), \alpha\subseteq\beta$ or $\beta\subseteq\alpha$
c) Total reflexivity: $\exists\alpha \in Ns\ (w)$ s. t. $w\in\alpha$
d) Local absoluteness: If $\alpha \in Ns\ (w)$ and $x \in \alpha$, then $Ns(w) = Ns(x)$
e) Closure under intersection: If $S \subseteq Ns(w)$ and $S\neq 0$, then $\cap\ S \in S$

The non-emptiness condition can be interpreted as seriality from Kripke frames, stating that each world has at least one set of neighbourhood-related worlds via the truth set instrument. Nesting gives us linear ordering, which secures the translatability of Kripke frames into neighbourhood frames, as the accessibility relation in Kripke frames is linear. Total reflexivity secures the condition of each world belonging to its own neighbourhood. That implies that the truth of some φ is related to φ obtaining in the world from which we defined the neighbourhood. Local absoluteness states that the neighbourhoods of a world w and any element that belongs to one of the proposed neighbourhoods coincide, i.e. are connected via the neighbourhood function. The condition of being closed under the intersection states that every subset of a neighbourhood, while the a) condition of non-emptiness obtains, itself constitutes a neighbourhood of the starting world.

Now it might be useful to show how Kripke models and neighbourhood models are related. We have seen that they effectively differ in the operational part of the model, the accessibility, or respectively, the membership relation of the model, but it is possible to show how they might be interdefined. I have stated that the neighbourhood models are a generalisation of Kripke models as they are capable of describing structures that Kripke models are not because of their non-linear behaviour. However, every neighbourhood structure is transformable into a Kripke structure by adding two formal conditions; (1) the model being closed on the superset of neighbourhoods, and (2) the model being closed on the intersection of its elements. The term that is used for a neighbourhood model that validates both conditions is *augmented neighbourhood*.

Pacuit (2013) displays a proof that each augmented neighbourhood is translatable to a Kripke structure and vice versa. This means that we are in a position to construct a model comprising two distinct accessibility relations, one of which is supported by a neighbourhood structure, and another that is supported by a Kripke structure because of the very fact that each neighbourhood structure that is shown to conform to the conditions of being closed on the superset of elements and being closed on the intersection of the elements denotes the same formal object – a Kripke structure.

As we have seen, CDL is calibrated to two distinct relations, one for hard, and the other for soft information. The hard information pertains to the set of epistemic cells, i.e. the list of possible situations that are compatible with everything that is known by an agent at some point in time, i.e. given situation. The logic for hard information in defined with a universal accessibility relation and is therefore supported by a Kripke structure. Here ought to be stated that, while CDL system's hard information accessibility relation is defined as an S5-supported relation, we might be keen on weakening it in order to become congruent with the epistemic theory that we have so far established. This would mean that the accessibility relation for epistemically defined indistinguishability is basically an S4.2 relation. I will discuss this in detail when displaying the completed model in the penultimate chapter of the dissertation when dealing with adapting the logical model to

Distributed Systems interpretation which will be calibrated to the systems presuppositions such as multi-agent interaction.

The soft information, conversely to classical factive information, ought to be given a different approach, as it does not conclusive establish a demarcational exclusion criterion between the worlds that are compatible with what is known or believed, but displays a sort of an ordering between the epistemically accessible worlds through the notion of plausibility. This relation appears to require a bit more structure than the Kripke frames are capable of providing, so we introduce neighbourhood frames to do the job of establishing ordering relations necessary for explicating the plausibility relation. We often use terminology not dissimilar to the one of soft relation in CDL in our everyday practises, where we deem a possibility more plausible than another. Let us consider for instance the following example;

Jane has arranged to meet a friend for a drink. The friends has confirmed that she was planning to come. Jane notices that her friend is late. The more time passes, she might consider different explanations of why her friend is late. When her friend was late 5 minutes, Jane was probably convinced that she might have been looking for a parking space, or that her bus didn't arrive on time. When 20 minutes passed, she might have considered different options such as that she might have had an accident or had fallen asleep before the time of their arranged meeting. Each of these is positively compatible with what Jane knows at a certain point in time. No option is considered by Jane that contradicts anything she knows. If asked at each moment, Jane is quite able to provide a list of situations that she deems possible and put them in an order of what she considers the most plausible and what she considers least plausible. For instance, the situation (or a possible world) in which her friend was abducted by a spacecraft of purple pygmy aliens is probably near the bottom of the list of plausible explanations. From this example it is quite easy to see how we can describe the logical structure of these two accessibility relations; firstly, it is quite clear that two epistemic situations must be epistemically equivalent in order to entertain the respective levels of their plausibility. Thus, we can say that in order for two situations be comparable in order of plausibility, they must not be epistemically distinguishable, i.e. they both must be compatible with everything that is known at a certain point in time.

I feel as if the reader is owed an explanation as to why exactly we have assumed two distinct formal structures (Kripke frames and neighbourhood frames) over the relations of accessibility and membership that would calibrate our models for plausibility ordering. The reason is that, while the partial-preorder of the S4.2 relation fails to account for the structural properties of the plausibility orderings, the plausibility orderings insist on comparing each pair of R-related epistemic cells. That means that all the situations that are encompassed in our analysis must be comparable to each other, notwithstanding the mutual linear accessibility. This guarantees that we are not only comparing the epistemic situations based on true facts about the system, but each fathomable *possible* setup within the provided framework. Such an instrument can allow us to express epistemic and doxastic update not necessarily through a standard dynamic logic, but based on plausibility relations between various factive and non-factive content.

DEFINING CONDITIONAL BELIEF

As we have seen, CDL stands for Conditional Doxastic Logic, so we ought to define what exactly is conditional-oriented about our belief system. The basic idea is that everything that we believe is based on a set of given body of statements that are *provisionally* taken to be true in some epistemic context. We can, thus, say that we believe some proposition B based on everything provisionally given so far. And although this instrument gives us the possibility of considering even the non-factive doxastic options, i.e. situations, I believe that this is the most natural way to express the idea of beliefs being based on any set of propositions that can be taken as evidential grounds for some belief. This virtually means that we can interpret it as establishing a function from some non-empty set of propositions to another non-empty set of propositions.

Let us first and foremost define the language that CDL uses in Backus-Naur form, after which we can lay out the axioms and inference rules of the system, before formally defining the notion of conditional belief supported by a neighbourhood frame.

DEFINITION (LANGUAGE OF CONDITIONAL DOXASTIC LOGIC)

$$A := P \mid \bot \mid \neg A \mid A \wedge A \mid A \vee A \mid A \supset A \mid Bel_s\,(A|A)$$

As the reader can notice, the system's atoms are propositional variables and the primitive negation ⊥ is introduced. With respect to operators' arity, there is one unary operator – negation ¬, and four binary operators – conjunction, disjunction, implication, and belief. All are extensional except for Bel_i, making it the only expressible modality within the system.

With respect to the rules that the system uses, we define the following two:

If $\vdash B$, then $\vdash Bel_s\,(B|A)$ (the rule of epistemisation)

If $\vdash A \supset\subset B$, then $\vdash Bel_s\,(C|A) \supset\subset Bel_s\,(C|B)$ (the rule of logical equivalence)

The former is obviously an instance of necessitation, securing that logical and mathematical truths be known by the agents of the system. This is obviously an idealisation, albeit an essential one, as the uniformity of behaviour of the system would not be otherwise expressible. The second rule is an instance of logical equivalence, sometimes referred to as the rule of material equivalence, and often introduced axiomatically, rather than as an inferential rule.

Further, we define the operator Bel_s in the following manner;

$M, w \Vdash Bel_s(B|A)$ iff $\forall \alpha \in N_s(w)(\alpha \cap [\![A]\!] = 0$ or

$\exists \beta \in N_s(w)(\beta \cap [\![A]\!] \neq 0$ and $\beta \cap [\![A]\!] \subseteq [\![B]\!])$

This basically means that the world w in the model M forces the belief of B under the supposition of some body of information A if and only if either of the following conditions are met; (1) the condition vacuously satisfied as the intersection of α and the truth set of A is empty, or (2) if there exists such element of neighbourhood in which the intersection is a non-empty set, and the element of that set which is the member of the truth set of A is also a member of the truth set of B. This is the case because of the behaviour of material implication (which is obviously the implication in use in Conditional Doxastic Logic); If the antecedent A cannot be satisfied by means of being a member of the intersection of a neighbourhood α and the truth set of A (the set of worlds that belongs to the world w's neighbourhood determined by the truth of the formula A), then the implication is true by definition. That means that no condition exists upon which B should be accepted, thus, the implication holds. On the other hand, if there is such a condition to be met for the formula B to obtain, then both A and B ought to be true, as that is the only possible interpretation for the conditional being true when the antecedent is satisfiable.

Now that we have a clear definition of what conditional belief is and when it is satisfied, we can take a closer look at the axiomatisation of CDL. As I have commented at the beginning of this chapter, CDL system behaves like a normal logic in the sense that it validates all the tautologies of Classical logic, alongside with the axiom of distribution and the rule of necessitation, but because of its internal conditional-based infrastructure it might be more natural to talk about it in the terminology of neighbourhood semantics. I have briefly mentioned that CDL validates two distinct accessibility relations, one for hard information, and the other for soft. The accessibility relation for hard information is an equivalence relation that is an S5 relation. Conversely, the soft accessibility relation (which we will mostly address in neighbourhood frame vocabulary) can be viewed to correspond to a KD45-modelled relation in Kripke frames, as it validates the axioms completely analogous to it, the only difference being the definition of a minimal well-formed formula as a conditional structure of the sort (B|A). So, as I have stated, we can consider the system normal as it meets the sufficient conditions for normality, but we still opt for defining it in neighbourhood frames that are canonically used for non-normal systems because it fits the conceptual groundwork of the system. As we have seen so far, for the construction of the model for this dissertation I have opted for an S4.2 validating model for knowledge, a CDL validating model for belief, and as we will observe in the following chapter, an S5 validating model for defining meaningfulness intrasystematically (as a separate relation) as the proposed structure should be calibrated for modelling epistemologies akin to verificationism.

Taking all this into account, and as the title of this dissertation suggests, all of the structures in the model are normal, which means that we have successfully preserved all of the semantic and syntactic properties that each normal system (from minimal K onward) possesses. It might be argued that it is a positive upshot of attempting to devise a normal epistemic and doxastic model. In order to remind the reader, I will return to this point in the final chapter of this dissertation when I discuss the metalogic of the proposed model. Now I feel it would be adequate to move onto the axiomatisation of CDL, through which we can observe its equivalence to a KD45 system with a differently defined basic structures.

DEFINITION.

THE AXIOMATISATION OF CONDITIONAL DOXASTIC LOGIC

The axiomatisation of CDL validates all tautologies of Classical propositional logic.

Ax 1 $(Bel_s(B|A) \land Bel_s(B \supset C|A)) \supset Bel_s(C|A)$ (distribution)

Ax 2 $Bel_s(A|A)$ (success)

Ax 3 $Bel_s(B|A) \supset (Bel_s(C|A \land B) \supset\subset Bel_s(B \supset C|A)$ (minimal change)

Ax 4 $Bel_s(B|A) \supset Bel_s(\neg Bel_s(B|A)|C)$ (minimal change 2)

Ax 5 $A \supset \neg Bel_s(\bot|A)$ (consistency)

Distribution is a standard axiom of normal systems and because of the basic structures of CDL, here it plays a role of modus ponens. It basically states that if we were to accept B under the condition of accepting A, and we were to accept B implies C under the condition of accepting A, then we would be justified to accept C under the condition of accepting A. So, we could say that it denotes closure of conditional belief under modus ponens. The following entry is the axiom of success, which corresponds to one of the postulates of AGM theory. It basically states that when some information is given to an agent as true, then it is believed by the agent to whom it was given. The following two axioms state that when a belief set is revised by some new body of evidence, the change within the belief set should be as minimal as possible. In other words, the agents should maintain as much of their belief set once a change happened. The last axiom simply states that the truth of any formula guarantees that its negation is disbelieved. This stems from the fact that we can rewrite $\neg A$ as $A \supset \bot$ when using primitive negation (bottom), and from the fact that basic structures of CDL are defined as conditionals, e.g. B|A is understood as if A then B.

Finally, just a few words on dealing with the complex infrastructure of CDL system. We have stated that it possesses two accessibility relations, one for hard and the other for soft information. In order to understand what that formally entails, we ought to be aware that each axiom within the system provides a restriction on the accessibility relation in both Kripke and neighbourhood frames. If this is the case, the accessibility relations (or membership, depending on which framing we are using) are defined through basic claims of the system.

In order to make sense of plausibility models, we might want to attempt to offer a natural philosophical reading of them. We have already observed some examples, but it appears that we need a bigger picture in order to get to the bottom of what they do within an epistemic-doxastic model. As the name suggests, plausibility models produce a sort of an ordering of worlds within the model by how much does an agent ascertain their probability given some pre-existing body of information. In other words, agents compare worlds, or rather sets of worlds, given everything they know by the projected level of their probability. Thus, the body of information that is relevant for constructing a plausibility model is basically the set of everything that is known or believed at a certain point in time. A natural reading of them would probably be in the vein of applying abductive reasoning in a specific epistemic and doxastic situation.

Let us take another example. Suppose I have bought a ticket for a bus on a relation Rijeka – Maribor. I come to the bus station and find the platform from which the bus is stated to depart. I see a bus parked there and check the writing on a cardboard in the front of the driver's seat which reads Rijeka – Maribor. Now it seems I am almost conclusively convinced that I have found the right bus. I enter it and wait for departure. At that point in time, given everything I know and believe thus far, I have formed expectations. I obviously expect the bus to depart at the stated time, and connect to the road for Zagreb, as that is the prearranged route in order to get to Maribor. But I do not a priori completely dismiss every other possibility taking place. Even more so, just as in the previous example, it appears that I can state that some are more probable or plausible than others.

The quite reasonable possibility, which might not be my first choice for the most plausible world, is one in which the bus departs later than it is stated on my ticket. Such a world still seems very plausible, maybe even more plausible than the first if I am a pessimist. Another world which is maybe further down the line of plausibility is the one in which the bus starts driving towards Ljubljana or Pula. Even more so, there exists a world in which the bus drives off into the port of Rijeka into the sea, and on the outer rims of plausibility is the world in which it flies off into the night sky. All of these worlds appear to be compatible with what I know so far in terms of their shared histories, and hence, my epistemic states, but I certainly have more reasons to believe that some of them have more or less merit for candidacy for being the actual world.

The plausibility models are what makes this kinds of comparison and ordering of the possible worlds formally expressible. We want our system to be able to state that some worlds are at least as plausible as any other in order to capture the essence of the ordering relations. The plausibility ordering relations are usually closed under reflexivity, transitivity, and connectedness, which at first seems at odds with what we have so far observed because for a frame to be closed under reflexivity, it needs to validate the factivity axiom (T axiom), which we obviously do not want for our belief theory to do. This is because we want to be able to state within the system that some agent has false beliefs, however, when we talk about reflexivity in the realm of plausibility, it is a much more palatable consequence. It is harmless for the system as it simply states that each world is at least as plausible as itself, which validates the notion of world identity in this kind of formal enterprise.

Furthermore, it also seems that some of the worlds that we are ordering appear equally plausible, or to use a technical term – equiplausible. So in order to arrive at a more natural reading of plausibility frames, we turn to the other side of the same theoretical coin- we display the system in the theoretical framework of Grove spheres. Grove spheres are essentially plausibility models based on the notion of world nesting. They are geometric representations of plausibility world orderings that are defined as a structure of concentric spheres. This gives the plausibility models a topological reading, rendering the model calibrated for translation into the Neighbourhood Models.

The spheres work as following; the innermost (central) sphere is the set of worlds that are R_k-accessible, in other words, the set of subjectively indistinguishable worlds that *must* be compatible with everything that the agent knows in some epistemic situation (or at a certain point in time, provided that time is formally expressible in the model). Once such a set is established, we are introducing the notion of epistemic cells. Epistemic cells are abstract states that are mutually comparable within the model. At this point, all the worlds within the central sphere are equiplausible, or of equal plausibility, by definition. This is the case because we have not introduced any variable that would make them discriminable for the agent. This will remain so if we feed only hard information into the system – all of the worlds will remain equiplausible to the agent in relation to what is known, which is the central idea behind hard information. Only when soft information is introduced, the type that does not constitute knowledge, but only gives an indication as to which world is more *likely* to be the actual world, does the agent start considering an ordering between worlds. They are now geometrically represented as concentric spheres that are encasing the central R_k-defined sphere. The smallest sphere, let us call it R_{b1}, that encases the central sphere comprises the worlds that are as plausible as they can get without being R_k-defined. In other words, all of the worlds that belong to R_{b1} sphere are (1) mutually equiplausible, and (2) at least as plausible as any other world belonging to a sphere, let us call it R_{b2} that encases the R_{b1} sphere. This structure can be continually developed in this manner until we run out of worlds that are in need of assessing in terms of their plausibility.

The notion of world nesting is in itself very simple – in each concentric sphere we nest worlds by equiplausibility. In other words, we can state that each epistemic cell is mutually comparable and that there exist a sorting of the cells into the concentric spheres that represent their perceived plausibility for the agent in question. Notice that the plausibility ordering is necessarily agent-relative, because of the obvious fact that agents do not necessarily share knowledge – even more so, it would be very unusual if there were to exists two agents with the same corpus of knowledge.

The following is a formal representation of epistemic plausibility models that support the structural inventory of the theory;

DEFINITION.

$$M = \{W, \sim_{s\in A}, \leq_{s\in A}, [\![\ ]\!]\}$$

The model is quite similar to each we have thus far observed, where W is a set of possible worlds, or the domain of the system, and $[\![\ ]\!]$ is a valuation function, supplying the semantics of the model. The accessibility relations that we have already informally captured are $\sim_{s\in A}$, the subjective indistinguishability relation for the central sphere of the system, i.e. the corpus of knowledge, and $\leq_{s\in A}$, the well-founded pre-ordering relation for the mutually camparable nested worlds of the system.

The proposed relations of the system, thus, are describable in terms of the following properties;

The first states that *plausibility ordering of worlds implies possibility*; If two worlds are $\leq_{s\in A}$-related, then they are also $\sim_{s\in A}$-related, as one cannot compare worlds in terms of plausibility if they are not epistemically equivalent for the agent. For this to make sense to the reader, we can observe these two relations as if one were to represent knowledge and the other belief – In order for us to be able to ascertain which of the two situations that I am considering in terms of forming a new belief is the more plausible, the two first have to both be compatible with what I know, i.e. they should rely on the same corpus of given facts. If that weren't the case, then the comparison would appear to be null and void, as the presuppositions of their potential introduction into the belief set would be irreconcilable.

The second property is the one of local connectedness – if two states, w and x, constitute epistemically indistinguishable states, then either $w \leq_s x$ or $x \leq_s w$. Either of the two options *must* be the case – either x is at least as plausible as w, or w is at least as plausible as x. The reader can notice that this dilemma still maintains that the two may be equiplausible, but it only states that if they weren't, then either one or the other would necessarily be *more* plausible.

This guarantees linearity of ordering, necessary for the system to be supported by a Kripke frame. However, as I have already indicated at the beginning of the chapter, most authors are keen on using neighbourhood frame vocabulary for speaking about Grove spheres and epistemic plausibility orderings, as the set-theoretical language better captures the way we perceive nesting of the worlds within the spheres of plausibility. If we are keen on constructing a neighbourhood model for the weaker accessibility relation (as it is the relevant one for this vocabulary of choice), we can express it in a following manner;

DEFINITION. $M = \{W, N_{i \in A}, [\![\ ]\!]\}$.

The accessibility relation, or rather a membership relation, that is displayed in this model covers the plausibility relation for soft information as it models full ordering of the worlds that are R_k related. In other words, the accessibility relation that captured subjective indistinguishability was an S5 relation, which is obviously supported by a Kripke frame, as its axioms validate a full equivalence relation between the worlds – each world has access to each. We can then compare the epistemic cells based on soft information which will provide the basis for world ordering based on their perceived plausibility in relation to the agent within the system. So far, this model validates mono-agent knowledge and dynamic belief, so we are still in need of some more structure in order for this formal theory to be calibrated to Distributed Systems Models, which will finally be displayed as dynamic complex formal structures that are capable of expressing both epistemic update and belief revision for multi-agent contexts.

So, as we have seen, the CDL system and its class of logics that deals with agent-relative plausibility defined through a conditional structure of the type A/A is validated by plausibility frames that can be translated into neighbourhood frames without any issues. Just as a remined at this point, I have previously stated that the Grove spheres and plausibility models are deductively equivalent, so by extension it follows that even if we were to frame our formal theory of plausibility in Grove's spheres, the neighbourhood models are calibrated to display this structure with the same ease. Furthermore, once we have displayed the structure in the neighbourhood models, we are quite close to being able to show that the system is supported by a Kripke frame. The elements that remain to be introduced in order to show this equivalence are the Axiom of Distribution and the rule of Necessitation. Once we establish that the system we are observing validates both, we are certain that the model that the Neighbourhood model is *augmented*. The notion of augmentation is defined within the neighbourhood models as validating closure under the intersection and closure under subset. We now know that the observed Neighbourhood structure validates both, so we can infer that it is augmented. Furthermore, we also know that if the Neighbourhood model is augmented, there exists a Kripke frame which validates the same structural inventory of our formal theory. Thus, this model is supported by a Kripke frame, rendering it a *normal doxastic logic*. This falls in line with one of the central concepts of this dissertation, which is to show that we can define epistemic and doxastic models that are dynamic, but still maintain normality. I will come back to this point within the seventh chapter of the dissertation that deals with modelling verificationism using this formal theory, but it is of importance to note this here as well, as the doxastic frame of Conditional Doxastic Logic we have been so far observing was the most contentious, as it is quite common (as in the case with DEL system) that adding dynamics to a static doxastic theory falls into non-normality, which would be in tension with the very title of this dissertation.

In the next chapter I will attempt to develop a base for a verificationist theory that should be supported by the formal theory we have so far been observing.

CHAPTER VI – VERIFICATIONISM

The plan for this dissertation was never to develop a fully explicated verificationist theory, as that would be far outside the scope of this project. The verificationist account that I commit to in this work is solely based on the idea of laying a cornerstone of a combined semantic and epistemic theory that would be able capture the notion of meaningfulness in modal terms, while maintaining the view that no statement that is unconceptualisible can play a meaningful role in one's cognitive economy.

As an upshot of this remark, the epistemic-doxastic medial structure of this model is by no means *necessarily* calibrated to a specific verificationist enterprise, but to any epistemological theoretical framework that (1) is not in tension with the postulates that were defined within the proposed logic of knowledge and belief, and (2) takes into account the limitations of the formal vocabulary that the model is capable of expressing. This being said, the *reading* of a formalism is always up for interpretation, because of the very fact that it is not content-laden. The things always boil down to the degree of its applicability to a certain set of theoretical statements that might comprise a complete philosophical theory, or just a provisional theoretical framework.

Verificationism itself is an antirealist epistemic theory that underwent many revisions for the better part of the last 110 years. At its core, at least in my view (that I will attempt to cover within the margins of this dissertation), lies a strict semantic groundwork validating an epistemology. What that means is that the main philosophical concern that verificationism addresses is establishing a set of epistemic criteria for defining what is knowable that should in some way correspond to a set of criteria for what is meaningful within a discourse. Meaningfulness is then established on the basis of understanding the conditions that *make* a statement true.

The main proponents of this epistemology were the members of the Vienna Circle, an informal group of individuals that shared an interest into developing an epistemic and semantic theory which would be strongly oriented towards somewhat radical empiricism and the study of science through the lens of formal logic. The term 'scientism' was often applied to the theories of the Vienna Circle, most of the times derogatorily, as the authors strived toward a scientific explanation that was based on unmediated observable experience. They relied on such a conception of experience, as they were of opinion that *that* was the only way in which we could coherently explore and systematise our environments.

A non-exhaustive list of the proponents of the theory of verification comprises the authors such as Carnap, Schlick, Neurath, Frank, Hahn, and Ayer. Even though their philosophy was quite revisionistic in any reasonable sense of the word, some groundwork was done before them by Ludwig Wittgenstein in his seminal monograph *Tractaus Logico-Philosophicus* (1921)[8]

In this chapter of the dissertation I intend to briefly touch upon some of the classical formulations of the verificationist theses, which should work as an anchoring foundation of what I perceieve as attainable and useful part of the verificationist enterprise. The key points that I will address do not tie in strongly to the classical formulations of such theories from the 1920s, but more to the contemporary variations that use modal analysis and intensional concepts to capture the relationship between meaningfulness, knowability, and cognitive conditions. Even with this being the case, I will still argue that the original members of the Vienna Circle got many things right. Even more so, some positions and arguments that they saw as non-salvagable under the attacks from the opposition can be adapted to contemporary frameworks and terminology and can often be circumvented or incorporated into the theory in one rendition or another. Furthermore, I will attempt to address the issue of *a priori* and *a posteriori* distrinction with modal analysis, which should clarify the picture of the source of information further. This will be proven to be especially pertinent to the Distributed Systems Models analysis of the Verificationist framework, as I will attempt to show that this distinction will determine the type of information that is either present or fed into the system. The *information type* is a crucial component of the DSMs analysis in terms of epistemic modelling, especially when we are dealing with a broader intended domain (as opposed to the one capturing exclusively highly idealised agents). In this chapter I intend to give only an indication of how I intend to specify the DSMs framework for modelling a verificationist epistemology, but the details of the case will be fully developed and explicated in the chapter VII. Apart from the DSMs' framing of verificationist theoretical structure, I will attempt to show that the language of the theory can be adequately translated and possess quite a natural reading.

Finally, in the last chapter I will continue the discussion about the question of why perceiving verificationism in a multi-agent system makes most sense. The application of Distributed Systems Models to the intended domain of non-ideal agents in an antirealist setting will determine the way we frame our conceptions of knowledge and belief, which is based on the logical systems we have so far observed; S4.2, CDL, and the algorithmic interpretation of S4.2. It should also be said that this interpretation is non non-ideal agent-specific. It can account for the behaviour of idealised agents as well, however, some aspects of the theory will not be operative when dealing with them. For instance, the algorithmic part of the equation of formally capturing knowledge will not have to be framed in a temporal dimension, as the idealised agents do not need time to process an

[8] Even though Wittgenstein himself dismissed the Vienna Circle's interpretation of *Tractatus*, it would be irresponsible to not to mention it as the spark of inspiration and the spiritus movens of the school of though that was later entitled 'logical positivism'. The term logical positivism nowadays is mostly viewed as either obsolete or even meaningless by some (as it was used in contexts that differed substantially), so we can opt for a more descriptive one, verificationism.

information and provide an output – their ’inference' is merely a set of logical relations. This being said, the agents of the system will also be framed in a verificationist setting in which their conceptions of meaning and knowledge are tightly connected.

SEMANTICS AND EPISTEMOLOGY

I hold a view that this interrelationship between meaning and knowledge is very natural in terms of our pre-theoretical assumptions when reasoning about our surroundings, anticipating some behaviour of the observed systems, and constructing explanations of why some event led to another. It should also be said that 'meaning' does not necessarily refer to linguistic meaning in this context. It pertains propositions, not sentences, so it refers to logical objects, not grammatical. The said interrelationship is here viewed as an essential part of framing of the notions of truth, meaningfulness, conceptualisation, and cognitive conditions. So, we can state in this context that meaningfulness is established on epistemological and logical grounds of defining cognitive conditions for a statement.

There have been many attacks on verificationism because of its very stringent conditions for meaningfulness, usually claiming that the verificationist conception of meaningfulness results in some form of epistemic chauvinism. However, there is a plethora of reasons why such an epistemology can help us build up a sound basis for what we are justified to believe. As one of the most prominent verificationists, A. J. Ayer is quoted in Misak (1995) to have stated; „The attitude of many philosophers reminds me of the relationship between Pip and Magwitch in Dicken's *Great Expectations*. They have lived of the money, but are ashamed to acknowledge its source.“

This position is in a way subjectivist, as it establishes truth as agent-dependent, but not agent-relative. That means that postulating an agent-based setting within the epistemic model is a necessary part of defining meaningfulness of the statements, as statements refer to what is congitively accessible to us, rather than what *is* the case in the world, as would be the case if we endorsed a realist epistemic enterprise. In other words, in a verificationist setting, there is little point in discussing the interrelationship between knowledge and belief as mere abstractions without introducing the bearers of such states. The other part of the equation, the position about statements not being agent-relative pertains to the presupposition that we establish common, non-private conceptual frameworks that serve as a basis for how we think and communicate about, as well as and organise the contents of our environments. Carnap (1950) in his seminal paper *Empiricism, Semantics, and Ontology* offers, at least in my view, the most succint and concise analysis of the foundations of verificationist understanding of the types of questions we can meaningfully ask; (1) internal, and (2) external questions of existence. Now, in order to discriminate between them, a bit of terminological and technical background is needed.

When we think about the sources of our prospective knowledge, one thing might be clear; either we know or believe some formula φ by being capable of infering to it via meaning or logical structure of it, or we might know or believe it beacuse we have empirical evidence for it. The former refers to analytical (semantic) truths , i.e. truths by definition and logical tautologies, while the latter refers to cases of being in some observational context that might indicate, confirm or verify the truth of the proposition in question. The first are, quite obviously, *a priori*, necessary truths, while the latter are *a posteriori* contingencies. This kind of convergence on coextensionality between a priori truths and analyticity on one hand, and a posteriori truths and contingency on the other is not uncommon in empiricist philosophical enterprises and can be seen in Carnap, Hume, Schlick, Hahn, etc.

In the interest of brevity, henceforth I will refer to these two *kinds* of semantic-epistemic relata as analytic and synthetic categories. But seeing that we are not directly exposed to propositions when interacting with our environements, we ought to first answer the question of what are we exposed to? In other words, which *kinds* of things can we observe. For instance, speaking as pre-theoretically as possible, the first thing that comes to mind are objects. Our minds (and maybe even senses) are capable, i.e. have a disposition of individuating discrete objects from our surroundings, and we have a language that is based on such a gestalt analysis. Secondly, we can observe properties. The hidden supposition within this theoretical foundation is that if an object or a property is inobservable, they are either reducible to the ones that are, or they play no role in our congitive economy. Thus, we can say things such as „I see a keyboard in front of me." or „This keyboard is black and the letters on it are white."

Establishing relations between objects and properties form propositions that should in some way describe and model the world around us. These two kinds of observable things do not comprise an exhaustive list of things that our language is capable of supporting, but we can take them as basic kinds of things that can serve as a solid basis for the construction of a conceptual system. This by no means implies that we cannot choose different constituents for constructing a framework, just that this basic example of doing so is simple enough to display how a framework can be understood and displayed within the context of its inner workings. Objects and properties are in this example, hence, just placeholders for constitutive elements of a conceptual framework.

As a notorious fact, Frege (1953), Carnap (1951) (and many others in the analytic philosophical tradition) notice the idea of individuating objects and properties is not reserved for the empirical realm of our epistemology, but is perfectly translatable and useful when applied to the mathematical and logical one. Discovering and systematising the common denominators in establishing a classification of useful conceptual and linguistic structures allow us to be clear about what we *mean* when we say something. But how are we to systematise them in a taxonomy that should be sufficiently expressive and adequately organised for our cognitive and linguistic needs? Carnap proposes the following: before even talking about concepts like objects or properties, we ought to establish a sort of a conceptual framework for each kind of things we are to coherently think or talk about. Such abstract entities bear several different names throughout Carnap's career, depending on the paper in which they were discussed, and consequently what he saw fit at a certain moment. The two most common terms were linguistic frameworks or conceptual frameworks. In this dissertation I will opt for the latter term, as I believe it better fits into the picture since our simple idealised processors are not necessarily language-dependant. They can possess conceptual content without ipso facto having a linguistic structure that relates to it.

So, in short, a conceptual framework comprises a set of presuppositions that are necessary for a concept to be fleshed out and precisely defined. In other words, we can understand it as a collection of conceptual requirements that define the nature of the thing we are talking about. We can talk about the conceptual framework of physical object, the framework of propositions, the framework of thing properties, the spatio-temporal coordinate system for phsyics, and so on. Each framework establishes the essentials for a concept to become operative and understood. Thus, in order to talk about objects such as numbers, we ought to introduce a framework that will support them. It will comprise a language, an axiomatic structure, a list of algebraic operations and their definitions, etc. One might say that we used numbers meaningfully long before devising formal devices such as axiomatic structures, but in the vein of what Carnap is introducing, we might say that a more rudimentary conceptual framework of numbers preceeded the well established formal apparatus that we possess nowadays.

So, the frameworks that Carnap introduces are in no way stabile and impersvious to change; as he states, we can unproblematically dismiss a certain framework if another one that is more practical and operative is found. As conceptual frameworks imply no ontological commitments, as they only pertain to ways we *talk* about and *conceptualise* things, they are easily dispensible, albeit necessary. Furthermore, if we were to use statements use concepts that are not supported by a singular framework, they ought to be supported by multiple frameworks' conceptual requirements in order to form a meaningful discourse. This approach generally guarantees that everything that is to be talked about is to be talked about meaningfully. The union of all chosen conceptual frameworks at a certain point in time for a group of agents can then function as a logical taxonomical structure of the natural language they use.

Although I share Carnap's sentiment from this article, I introduce this kind of theoretical foundation for a rather different reason; where he attempts to show that insisting on ontological (and by extension – semantic) parsimony when constructing a system of conceptual and linguistic forms is very much barking up the wrong philosophical tree, I want to establish that the inputs that our simple active processors acquire from the NULL processors[9] ought to be well-structured within the system, so that we can determine which

[9] Chapter 4, Algorithmic Knowledge and Distributed Systems Models

kinds of algorithms are appropriate for their problem-solving purposes. And here I feel there should be another point of clarification in what I mean by describing an input as well structured – the input must be system-dependent in a sense. If we were to take the input to be a natural number, our processor's local state (and its algorithmic partition) must be clabirated to process it in an appropriate manner.

NB Input into our simple processing units might not be simple – its conceptual content might be supported by several different conceptual frameworks. So if we were to understand our node as a simple question-answer machine, when we ask it „Is the chair that is in front of you now green?“, we must be certain that its local state's context partition is calibrated to the input's conceptual and logical structure – we must be certain that it contains the adequate presuppositions concerning physical objects (chair), spatio-temporal relations (in front of you and now), and properties and their mode of attribution (green for physical objects). So, we can ask it „Is the given prime number be red?“ No, as the conceptual framework for numbers, and eo ipso, the processor's context partition of the local state in question, is not calibrated for attributing colours to primes.

Our processors and their local states are perfectly capable of being programmed to form conceptual bridges between their respective frameworks – they can obviously deal with red chairs, 10 apples, or brooms in closets or behind doors. This is the case because the types of entities (physical objects, colours, spatial and temporal relations, etc.) introduced as inputs are not introduced into discourse under conflicting presuppositions. To say that a number has a colour is meaningless specifically *because* the conceptual frameworks for numbers are not calibrated for such properties.

Furthermore, for Carnap, the entirety of this apparatus serves the purpose of explicating which classes of questions about certain entities can be asked. Carnap differentiates between internal and external questions of existence. The former relates to intrasystematic questions (operating under the suppositions of the framework in question) about its elements, whatever they may be. So, a question „Is there a natural number larger than 4 and smaller than 6?“ is an intersystematic question because the proposed framework provides us with the tools necessary to answer it. Conversely, the latter type of questions does not pertain to extrasystematic questions on the *truth* of the proposed frameworks*, i.e. their correspondance to reality*, but should be viewed solely as a *study of application* of the proposed frameworks. If the introduction of some framework generated positive consequences in terms of our engagement with our surroundings, meaning that we might have better anticipated the future behaviour of the observed system, or that we had more concise and better systematised description or explication of it, then we might be keen on continue using the conceptual framework. If that proves not to be the case, we can easily discard it or substitute it for another theoretical framework that suits the purpose better.

This ties in rather well with the DSM approach to agents as processors and their systematic requirements. External questions of existance, in this case, might be understood as *ways in which we encode information from the NULL processor.* The canonical function of NULL processors has generally been viewed as injecting an input from the world (the external context) into the system of the interconnected nodes. But rarely has it been discussed which form does the input take. I attribute this as a symptom of the majority of theoreticians dealing with the subject being epistemic realists; in their view the input *is* the state of affairs in the world. The φ being fed into the system *is* the fact in the external world. But the situation is a bit more complex than that. We simply must venture into the question of how the information is codified and understood by the processors and how their local states are attuned to information received as input.

EXAMPLE ANALYSIS

As a short reminder of the question at the core of constructing a philosophical theory around a formalism that should be able to underpin it is *how to read* formal models, in our case models of knowledge and belief. I have discussed the infrastructure of epistemic and doxastic models in Kripke frames and stated that they comprise the set of possible worlds W, the binary accessibility relation R, and the valuation function ⟦ ⟧. Even though we have hopefully conclusively covered the role that these three components play within the formal aspect of the theory, I feel we should spend just a bit more time specifying what their philosophical interpretation should look like in order to fully grasp how we can be certain that no terminological tension arises when establishing whether one can work as a basis for the other.

The first, as we have seen in the second chapter of the dissertation, is the set of possible worlds W – a list of situations that we take into account when constructing a model. Philosophically speaking, the set comprises all sorts of situations, the ones that are compatible with the facts of our actual world, or our knowledge and belief, alongside ones that aren't compatible with either. In the set W we can find worlds where Tutankhamon, the boy pharaoh, was the guitar player for the Beatles, and mother Theresa founded the automobile company Aston Martin. We can say that the worlds that resemble ours in almost every relevant aspect are 'close' to ours, e.g. the world in which everything is the same as ours, except for the fact that I turned on my laptop with my right hand instead of my left. The 'further' worlds are the ones that are different from ours in some meaningful ways, e.g. the world in which there exists another colour within the light spectrum that is discernible to us, irreducible to any that we have so far observed.

So, how do we establish which of the possible worlds or situations matter to us as theorists in terms of modelling the world in any meaningful way? Well, the answer is, as always, depends. This relativity can be bound to what we, in fact, want to model. If we want to model a coherent depiction of the world that corresponds to the established facts of the world that we live in, then we are bound by the *known truths* of our surroundings. The criterion of which worlds constitute *viable options* is defined with the second component that was introduced, the binary accessibility relation R. It tells us which of the worlds are 'accessible' from our starting position in a way that the formal structure of the theory allows.

Such a conception ties in very naturally with the verificationist framework. Here I offer a not very common formulation of the verificationist thesis, but I think it suits the model rather well, while preserving the verificationist intuition about in principle testability and meaning. We say that a non-logical, i.e. empirical sentence is meaningful, in other words that it *expresses a proposition* iff (1) there exist at least two comparable worlds within a model, one in which it is true, and another in which it is false, and (2) we are capable of defining cognitive conditions for establishing in which of the two (or more) worlds we are actually in. As I have noted, the proposed model defines knowledge as pin-pointing the actual world within the set of epistemically and doxastically accessible ones. The cognitive conditions in the proposed definition are virtually reducible to modes of verification, i.e. ways to check if the said proposition obtains in some world or not. The term *verification* should not here be read in a traditional logical empiricist manner, as I do not claim that a set of observable statements logically entails a theoretical one, just that the outline of the principle of verification is explicable through the framework of possible worlds. A similar idea was put forward, although in quite a different manner, and for quite different philosophical reasons by Gordian Haas in his seminal work *Minimal Verificationism: On the Limits of Knowledge* (2015). Our logic for belief revision (CDL), which was displayed in the previous chapter of the dissertation, gives us a guarantee that all the epistemic cells within the model are comparable by the principle of local connection. This entails that each world that an agent considers possible under soft information is comparable to each.

Let us take some examples to clarify what is meant by this and how it affects our criteria for meaning. We can attempt to establish the verifiability conditions for the following five sentences:

(a) There are currently exactly four security guards in the Governor's palace in Rijeka.
(b) Diogenes' barrel was a meter and a half wide at the opening.
(c) The nine people who opened the Tutankhamon tomb died because of its curse.
(d) Free will exists.

The first sentence (a) at first appears the simplest in terms of analysis. Currently, we are unsure about whether it is true or not. However, in terms of what I know and believe, there are worlds that are accessible to me in which there are, in fact, four security guards on the premises of the Governor's palace in Rijeka, and there are those in which the number might be lower or higher. That basically means that we do possess the means to conceptualise coherent situations that differ in this one variable, which constitute the necessary and sufficient conditions for the sentence to be understood, i.e. meaningful. In fact, I can get out of my chair at this point, walk down to the bus station on Sveučilišna Avenija, get on the bus to the Governor's palace and see for myself if there are four security guards or if there are fewer or more than four.

The second sentence (b) is of a different sort when we think about it. I most certainly cannot take a plane to ancient Greece and measure the diameter of Diogenes' barrel. However, I *can* do something else; I can conceptualise the worlds in which the diameter was smaller and worlds where the diameter was bigger than one and a half meters. The fact is *in principle* verifiable, I just have no means to physically do it. But my very *understanding* of the means to verify it (even in principle) constitutes enough reason to deem the sentence meaningful, as my mode of conceptualisation of two or more situations that differ in this one variable renders it such. If I would not be able to even conceptualise the means of verification of this difference between the worlds, then it would appear to fail this test.

The third sentence (c) is a bit more peculiar from the other proposed examples. After the expedition in 1922, the people who opened the Tutankhamon tomb died in some odd years after the fact. It all started with Lord Carnarvon – in 1923 the media reported he had died from a mosquito bite, but his death kicked off rumours that his death was caused by the Pharaoh's curse for opening the tomb. His death was followed by deaths of other eight expedition members within the duration of the following 17 years. Each death was heavily publicised as the continuation of Tutankhamon's curse. Let us take a look at an analysis of this situation from our semantic-epistemological framework. Can such a claim be construed as meaningful?

The first thing we have to look at are the conditions of verification. How would we establish a meaningful contrast between this explanation of events that following the opening of the tomb from the one that the participants in the expedition died from various causes. Firstly, the media did not refute them dying from various causes such as mosquito bite related illnesses, heart attacks, suicides, or house fires. They just *added* the extra step in explaining the deaths of people who opened of the tomb. The claim was that their various causes of death were *propelled by the curse*. So, if we wanted a modal reconstruction of this example, we would ask ourselves to conceptualise two distinct possible worlds, one in which the expedition members died from various causes, and another in which they died of various causes that occurred *because* of the curse. And then we should attempt to determine what *test* we could in principle construct in order see which theory explains the events better.

Of course, no test is to be found, as we cannot, even in principle, determine the inner workings of a curse (whatever that might be), as no variable of the proposed theory is measurable, observable, or in any way cognitively accessible to us. This basically means that such a theory would be considered meaningless when adopting this flavour of verificationism.

The sentence (d) is also interesting when considered within this theoretical framework. The problem of free will is almost universally lauded as a substantial philosophical problem nowadays. In fact, it is one of the most prolific topics in philosophy, as it piques an interest in many theoretical philosophers. However, just as it was the case with the last example that we looked at, it falls short in terms of the analysis that I am proposing. When comparing a possible world in which we have free will with one in which we do not, it is quite unclear exactly *what difference between them* we are looking for. Although this problem appears quite intuitive to us, when we look at it for some time, the situations in which we have free will, whatever that would imply, are indiscernible to us from ones in which we do not on the basis of measuring each of the cognitively accessible variables. This in itself is sufficient for rendering the problem meaningless within the proposed framework. Such an analysis might stumble upon a substantial pushback from the current philosophical mainstream, however, understanding the conditions for meaningful statements is crucial for establishing a basis for setting the stage for the formal model that I am proposing in this dissertation.

Again, thanks to prof. Berčić excellent comment, here I see fit to introduce an example that he recalled that should clarify this matter further. Shortly after establishing the theory of general relativity, Einsten (1916) published a paper entitled Approximate Integration of the Field Equations of Gravitation in which he showed that general relativity anticipated the existence of ripples in spacetime, i.e. gravitational waves that propagated at c – the speed of light. At this point, he had no idea how we could observe their existence, and even less so – how to measure them. Because of this, in 1963, he published a second article entitled „Do gravitational waves exist?“, in which he stated that, as we have no evidence we could ever empirically capture the notion of gravitational waves, they might just be mathematical artifacts of the theory. In other words, he doubted that any *empirical content* can be meaningfully assigned to them, as we have no means of neither observing, nor measuring them. Speaking from the view that I hold in this dissertation, the worlds in which they 'exist' were subjectively indistinguishable from the worlds in which they do not. It was not until 1937 that he changed his mind, when he postulated their existence as physical phenomena (albeit without having a devised method of observing or measuring them.

But in the verificationist framework of this work we can ask „What was his theory *about* before we had empirical content for the theoretical assumption he was making?“. We can approach this problem from two distinct standpoints. Before feeding any empirical content into this currently empty notion, they were either (1) just a part of the uninterpreted calculus that the theory uses or (2) there always existed empirical content behind the notion of gravitational waves, but it was yet to be discovered.

The first option would render the statements about their existence empirically meaningless in the proposed framework of this dissertation. The calculus itself was unproblematic, but the word would simply *not denote* anything *at that specific point in time*. Only when we assigned empirical content to the terms by means of constructing tests (even in principle), do they start having a role in our cognitive economy, and hence become a meaningful part of our physical theory. The second option is, quite obviously, a realist one, which I do not conform to in terms of my position in this work.

As a reminder, the final chapter of this text should show how we can model an epistemic theory with a formalism that should underpin its statements with its axiomatic schemata. This discussion about applying the formalism I have selected to an epistemic theory will be discussed in detail further in the next chapter, but it appears that a short preview wouldn't hurt, as it might help the reader understand the purpose of this chapter. We defined a set of processing units (processors, agents, nodes) of the system that are interconnected, and their local states correspond to variables in their environment. We claimed they are not inert, which means they are capable of interacting with their environment, solving problems, and adapting their behaviour to the external variables, proposed as a form of an input from the NULL processors.

In what we have seen in the fourth chapter of this dissertation, their behaviour is quite in line with the basic verificationist concepts that I have here proposed. They are presented with an input, and after their epistemic update, they consider the possible situations that differ from one another in some substantial and measurable way and their consider the situations viable iff they are compatible with what they know or believe.

Now, after considering the proposed examples, in the following part of the chapter I will display the proposed criterion for the verificationist understanding of meaningfulness formally.

THE PRINCIPLE OF VERIFICATION

In order for our theory to become effectively calibrated to a verificationist epistemology, a few further developments ought to be made. As we have discussed so far, the epistemic-doxastic model that we have been investigating was essentially being presented as a basic structure that will determine the meaning of the our epistemic and doxastic terms. The structure itself was complex, as it validated two distinct accessibility relations, one of which was supported by a logical theory of knowledge, and the other by a logical theory of belief. These two were interconnected on both structural and axiomatic level, as we have managed to establish their interrelationship formally by introducing a bridge axiom between them. The bridge axiom itself did not work as a simple definition, as is the case in alethic systems, where $\Box$ was translatable to $\Diamond$ simpliciter, and vice versa, but it actually said something of substance about both. It defined belief through knowledge on a theoretical level, which means that it offered a sort of a formal counterpart of a definition within our epistemic theory. But in order for our model to essentially capture the framework of verificationism, it appears that more structure is needed.

As I have stated, the natural thing to do so far was to define the epistemic-doxastic model as a basic structure without a need for a substructure or a superstructure. But for our model to be interpreted as a verificationist one, we ought to add another layer. Not only that, but it appears that the structural layer that ought to be defined will take the original model's place as a basic structure, upon which we are able to define a superstructure that will formally determine our boundaries of knowledge and belief. Naturally, the basic structure will be the semantic screen that will establish the candidacy of our well formed formulae for constituting meaningful statements. This basically means that there are statements in our language that are carried by well-formed sentences which fail to express a proposition. This is so because for something to be a proposition, it ought to bear truth value, or to speak in Fregean terms, its extension must be either truth or falsity. It is clear that there are many statements in our natural language (and unfortunately sometimes scientific) that meet the syntactic criterion of being a well-formed formula, while failing to express anything that can be awarded truth value. This point, however, might be contentious. It must be stated that the formal definition of meaningfulness is not universally accepted. In other words, the definition I am proposing is domain-specific in the sense that it excludes a great deal logics which might support a broader array of statements in terms of considering them meaningful. So, in order to be clear, we must address the relationship between formalisms, i.e. logical systems, and natural language.

Firstly, as a notorious fact, it should be clear to the reader that a logic and its language are not the same. It should further be clear that now we are observing the language of logical systems, and are not taking into account the inferential apparatus or axiomatic statements of the system in question. One of the few things I should want to note here with regards to the logic itself (not its language), is that the system we are observing in a modal extension of the classical propositional account. Now, with respect to the language that a logic uses, it is formally restricted to a number of logical constants and extralogical variables. It is canonically presented in the Backus-Naur form, which concisely defines what constitutes a well formed formula for the system in question. This allows the system to form complex sentences and evaluate them by means of the proposed semantics. But, here we arrive at the problem of formalisation. The phrase 'indeterminacy of translation' has a long and rich history in the annals of analytic philosophy, but instead of the Davidson-Quine kind, we are dealing with another. When we deal with translating (or maybe rather representing) a statement of a natural language into a formal one, what tacit presuppositions are we taking with us? I hope that the reader agrees with me that there exist statements in (any) natural language that bear no meaning, i.e. do not refer to any state of affairs in any of the possible worlds. I have addressed some in the chapter on verificationist epistemology. Once such meaningless statements are said to be represented by some variable within the logical system, we are in some way claiming that they in fact possess meaning by means of attributing them some truth value. So, it appears to be necessary to propose some kind of candidacy criterion in order to ascertain which statements are capable of expressing states of affairs in the world. Otherwise we would be bound to the view that all syntactically well-formed formulae talk about the world in some meaningful way. So now we have to be clear about how we approach formalisation at all.

The criterion that I here propose, hence, is not an internal instrument of the epistemic-doxastic logic that we opt for in our model, but should rather be viewed as a metastatement concerning concepts and their prospective properties when we attempt to formalise a natural language discourse. If we were to view this from a structuralist perspective (and I think we should as it can be proven to be useful), we can observe the proposed criterion as a sort of a layering of the formal theory.

As we have seen in chapter 6, the proposed line of inquiry when establishing a basis for conceptual analysis of objects, properties, and their relations corresponds rather well to Carnap's paper Empiricism, Semantics and Ontology (1950). In order to briefly remind the reader, the conceptual infrastructure of the proposed classes entities is based on a set of presuppositions that determine the categories of objects that we want to speak about. As Carnap thoroughly discussed, the introduction of a class of objects does not in any way entail ontological commitments, but only provides a conceptual instrument of clarifying the ways we choose to speak about the world.

As far as the few remaining strictly logical notions enter the discussion, we can state that, as a consequence of using the classical account, our objects cannot, for instance, be attributed some property and not be attributed the same property in the equivalent epistemic-doxastic situation. As my epistemology is antirealist, I choose not to speak about if real-world object are capable of behaving in a manner that contradicts this theoretical assumption. Thus, by choosing this metaframework, we are in a way dismissing logics such as free logic, paraconsistent and paracomplete logics. Nota bene, this decision is only in virtue of constructing a model that is capable of supporting the epistemological presuppositions here presented, and in no way a claim about the legitimacy of such systems outside this discussion. While I attempted not to make this the longest disclaimer in the history of analytic philosophy, a single further note should be made that the formal metacriterion of meaningfulness that I am here proposing is not bound to one specific epistemology, but is capable of supporting a wide spectrum of theories and systems that would validate it.

An upshot of choosing this criterion is such that meaningfulness is not often immediately determinable. Sometimes we will operate under the assumption that a statement is meaningful, and only when we attempt to reconstruct its logical structure will we discover that nothing of substance lies behind it. Mutatis mutandis, a statement might be proven to be meaningful upon learning some new information about the world. For instance, technological advancements made it possible for us to observe phenomena which we previously could not, and thus a statement regarding such phenomena, along with the theoretical presuppositions that we needed to introduce to deal with it in a clear manner could be proclaimed to be meaningful only upon its introduction to our experience and discourse.

So, in order to define this condition of meaningfulness both theoretically and formally via the apparatus of modal systems, we might say the following; A statement is meaningful if there exist at least two mutually comparable and up to that point subjectively indistinguishable worlds, one in which the statement is true, and another in which it is false. Understanding a statement is further restricted by producing cognitively operative conditions (with respect to the statement in question) that must be met in order for it to have a function in our cognitive economy. This means that we must be able to determine under which conditions does the statement obtain, and under which it doesn't. Therefore, I propose that the model for verifiable, i.e. meaningful statements be defined as following;

PRELIMINARIES. $M = \{W, \sim_t, [\![\ ]\!], T\}$, $T = \{t_0, t_1, t_2, ..., t_n, ...\}$

DEFINITION.

$s \in \mathscr{L}_v (t_1)$ iff $\exists P \subseteq W, f(s) = P, \exists w, v \in W$:

(1) $w \sim_{t_0} v$,

(2) $\forall t' (t_0 < t' < t_1 \Rightarrow w \sim_{t'} v)$,

(3) $w \nsim_{t_1} v$,

(4) $M, w, t_1 \vDash \varphi$ and $M, v, t_1 \vDash \neg\varphi$.

A proper subset $\mathscr{L}_v$ of the set $\mathscr{L}$ of wffs is a set of meaningful statements, i.e. a statement s is meaningful at t_1 iff there exists a state of affairs (a proposition, or rather a set of worlds) to which the statement is associated by a mapping function f. For this to be the case, we postulate two worlds w and v (playing the role of partitions in the model) which are subjective indiscernible to the agent at some point in time t_0. They only become discernible to the agent at some point t_1 when he/she is able to construct, or discriminate between the situation in which the relevant statement obtains from one in which it does not. This is in function of constructing cognitive conditions for the statement, rendering it meaningful. The temporal element simply allows for the notion of meaningfulness to work akin to a scientific discovery – once we have devised adequate conceptual or technological tools for examining and testing (even in principle) the relevant variable are we able to make sense of the statement that addresses it. As the logic that supports is classical, tertium non datur is validated, thus the system is decidable. This basically means that, when faced with a dilemma, disproving a negative is sufficient to establish the affirmative.

So, we are basically stating that for any statement to be meaningful, we ought to be able to establish cognitive conditions of discrimination between at least two situations in which it is true from ones in which it is not. The meaningfulness that is defined in this manner diverges quite a lot from the classical verificationist tradition, as it does not relate to de facto verification, but only in principle. To take one of the examples from the chapter on verificationism in this dissertation, we might never know in any sense (technical or pre-theoretic) what was the radius od Diogenes' barrel, but we know how to discriminate between the world in which it was one and a half meter or less from the world in which it was more that one and a half meter, or rather we can imagine how we would acquire such knowledge, given the circumstances of temporally co-existing with his barrel.

To be absolutely clear on the account of verificationism that I am here laying out, apart from it validating 'in principle' verification, it does the same thing for 'in principle falsification'. In my view, at least as far as I am willing to go with the theory of verification that I want to formalise within this mode, there is no substantive formal or conceptual difference between verification and falsification in principle. Both verification and falsification are based on a simple idea of developing cognitive conditions as a groundwork for establishing the semantic notion of meaningfulness. The only difference is the *type* of statements that they refer to. As it was clear to authors such as Carnap, Hempel, and Schlick in the 1940s, and as it is clear to contemporary theoreticians as Haas (2015), there exists a problem with verification in principle of general statements that operate under an (possibly) infinite domain, and there is a mirroring problem for falsificationism when it comes to falsifying existential statements. Hence, I think we are justified to conclude that the principles of verification and falsification are conceptually equivalent, as both presuppose the same structure of our experiential engagement with our surroundings, but they obviously *pertain to different kinds of statements*. Therefore, whatever logical theory we opt for when attempting to model some sort of verificationism, what we are doing is in fact modelling a disjunction;

φ is a meaningful statement iff it is either verifiable or falsifiable in principle. In other words, φ is a meaningful statement iff there exists a set of cognitive criteria under which we can construct a test of whether φ can be either verified or falsified. Each statement that fails to pass this test can be considered meaningless and cannot eo ipso be fed into the layer of the model that captures epistemic and doxastic states. This is because of the simple reason that meaningless linguistic structures (statements) can be neither known nor believed because they cannot be *understood*. If the reader minds my choice of terminology, I feel perfectly comfortable with changing the notion of understanding with one of conceptualisation, thus rendering the previous statements such that becomes about conceptual analysis.

The *'verification or falsification in principle'* idea has received some attention both in the golden years of verificationism and nowadays (Haas, 2015), however it was always dismissed for one reason or another. I will attempt to show that it works as the best rendition of the theory, and that it fits very well together with the formal criterion I have attempted to develop for this model.

FITCH'S PARADOX OF KNOWABILITY

The quarrel over the formulation of verificationism has plagued a great number of philosophers in the first part of the 20th century. Various attempts played with the notion of restricting the domain of knowable statements by introducing conditions of cognitive meaningfulness, and establishing a non-ambiguous interrelationship between semantics and epistemology as the basis of the system.

In 1963, in his paper "A Logical Analysis of Some Value Concepts", Frederic Fitch has formulated a theorem that he proved was a logical consequence of one of the possible definitional formulations of the theory of verificationism that caused an uproar in the academic circles. In it he appears to have shown that if we were to accept the statement that all truths are knowable in a verificationist framework, then by extension of the proof, all truths ought to be already known. His paper was almost universally lauded as a knock-down argument against any verificationist theory that would accept the proposed formulation, as no one would be keen on accepting the preposterous conclusion that all truths are known.

Formally, Fitch's formulation is displayed as following:

DEFINITION. $\varphi \rightarrow \Diamond K\varphi$.

From the formulation, a proof is offered that shows the following theorem:

THEOREM. $\varphi \rightarrow K\varphi$.

Although this is not a focal part of this dissertation, I think it serves solid ground for discussing what we are keen on accepting when defining a verificationist theory. Here I offer my reading of verificationism and what I believe went awry when Fitch introduced his formulation of the theoretical assumptions of verificationism that lead to the unpalatable conclusion he arrived at. As I have introduced at the beginning of this sub-chapter a model that defines the meaningfulness relation as a basic structure of verificationist theories upon which we can build a superstructure in the form of an epistemic-doxastic layer of formally representing an epistemology, I will rely on this structure when discussing Fitch's paradox.

So, in order to see how we can reinterpret Fitch's formulation of the problem, a few theoretical notions must be clarified and delineated. Those are: (1) verifiability in principle, (2) knowability de facto, (3) verification de facto, and (4) meaningfulness . Firstly, the semantic screen that I have introduced defines the verifiability in principle part of the equation. The screen provides a non-exhaustive list of statements that can be considered meaningful under the suppositions of this theoretical framework. As we have seen, it works through comparing two or more otherwise indistinguishable situations in order to establish cognitive conditions for meaningfulness of some proposition being true in one world, while being false in another. In other words, if we were in an epistemic situation in which we could de facto test in which world we actually are, then the statement can be considered meaningful. Thus, this displays the interrelationship between the first and the last notions that we have observed in this analysis.

The second part of the equation regards the second and the third notions that establish a relationship between some proposition being verified on on hand, and knowable de facto on the other. Returning to Diogenes' barrel example, it is quite clear that we are able to establish meaningfulness of the statements regarding its diameter as a variable that will allow us to distinguish between the world in which it is meter and a half or less from ones in which it is more than that. However, no verificationist in their right mind would claim that the statement is knowable de facto, as Diogenes' barrel no longer exists (and we do not even have at disposal the testimonial evidence regarding its diameter), so that statement, although meaningful in the context of stipulating the conditions of cognitive meaningfulness within the margins of this analysis, is most certainly not knowable de facto. This offers a clear line of distinction between statements that are meaningful from one that are knowable de facto. So, the interrelationship between verification de facto and knowability de facto is much less problematic aspect of this theoretical construction, as it follows a rather natural reading; When a proposition φ is verified, it is most certainly knowable de facto, and obviously even more than that, it is most certainly known if we operate under the assumption of our inferential system being sound.

So, finally we arrive at Fitch's proposed formulation of the verificationist Cartesian statement that we have seen formally states: $\varphi \rightarrow \Diamond K\varphi$. While I obviously have no issue with the provided proof, as it does what it set out to do, I believe that the basic statement does not correspond to the theoretical layering we have observed in this chapter. The natural way to read the statement that Fitch provided would probably be that meaningfulness of a statement implies knowability of it. However, as we have seen, this is most certainly not what we want from our theory. The in principle verifiability of the statement φ says virtually nothing of our capacity, albeit situationally determined, to know it. In other words, the meaningfulness of the statement about the Diogenes' barrel's diameter does not in any substantial way entail us being in an epistemic position to actually inspect whether it obtains or not. It only tells us that if we were in such an epistemic situation, we would be able to determine its truth value.

So, reading Fitch's formulation as 'each meaningful statement from the proper subset $\mathscr{L}_v$ is de facto knowable' is out of the question. We might want to inspect alternative readings of the statement at hand. The next one is based on a simple but relevant observation that it appears natural, when we use the English language as a metalanguage for reading formal statements, that we add the semantic element of 'being true' to our formulations. In other words, we would read out Fitch's statement as 'if φ were true, then it would be true that it is knowable'. As the 'truth' that constitutes the antecedent part of the conditional ought to be translated into verificationist terminology to account for 'conditions of verification', we would certainly be keen on understanding the statement φ of actually meeting the condition of verification, and thus we might want to read it as 'if φ were verified , then it would be knowable. I believe that this statement should work for everyone, including non-verificationist theoreticians. Furthermore, if we were to accept this reading, and followed the line of Fitch's proof, we would arrive at an equally plausible statement of any verified φ being already known.

Along the lines of this reading of Fitch's formulation, it might be useful to remind the reader of the simple fact that this statement is situated in intensional semantic space of a modal system. That means that when we formulate the antecedent of the conditional as 'if we were to accept φ', the natural question comes to mind; 'In which world'? Stating some proposition in intensional space warrant the explication of where we claim it obtains. So if we were to state φ obtains in the world w, by means of Fitch's theorem we simply state that within the world in which it obtains – in the verificationist terminology 'is verified', then it is knowable, and by extension of the proof known in the world in which it is verified.

The only remaining plausible reading of Fitch's paradox of knowability would be the one that states that 'each in principle verifiable statement is in principle knowable'. I hope to have shown that 'in principle knowability' is of no consequence in this analysis, as it does not have a natural reading apart from that of being meaningful. Knowability in principle basically refers to statements that are discriminable under cognitive conditions stipulated by the infrastructure of the semantic screen model that I have proposed at the beginning of this chapter.

CHAPTER VII – MODELLING VERIFICATIONISM IN DISTRIBUTED SYSTEMS MODELS

The following is the final substantive chapter of this dissertation and can be understood as a sort of a synthesis of the discussion on formal modelling of epistemic theories, with a special emphasis on modelling the theory of verificationism, as the model that I have been constructing validates such an epistemology. This chapter further comprises my 'central thesis' of this dissertation for a lack of a better phrase. As I have stated in the introductory chapter, the final product of my research is not a classical statement which represents a philosophical position, but a *formal structure* that supports a philosophical position. I have also stated that this dissertation does not discuss any logical system implemented in the construction of this model as *correct*, but only applicable for modelling some phenomena. This is quite obviously in vein with the verificationist spirit, as I hope the reader understands my motivation within the context for claiming that more than one theory (formal or not) can possess a high degree of application when attempting to explain, explicate, or anticipate the behaviour of a certain system. The formal structure that I have been addressing is a complex epistemic-doxastic model divided intro three layers. The following list should serve as a condensed and non-exhaustive explication of the layers, each of which will be separately addressed in their own part of the chapter;

(1) The base structure, *a semantic screen* that works as a function between the domain of a natural language $\mathscr{L}$, and its codomain $\mathscr{L}_V$, the language of meaningful statements within the verificationist framework. Only and all statements that 'pass the test' of the semantic screen are such that they can be fed into the second level of the structure, the epistemic-doxastic level. The screen will be formally defined through a *subjective indisinguishability* relation which will compare worlds on a newly established discrimination criterion, generatin a proposition, for each input from the natural language $\mathscr{L}$. It is supported by an S5 frame with a temporal component.

(2) The *epistemic-doxastic structure*, which clearly explicates the epistemic and doxastic postulates that this model validates through axiomatic schemata is established on two distinct accessibility relation. The first accessibility relation pertains to modelling knowledge, while the second models belief. Both are supported by a Kripke frame, however, the doxastic relation was in need of a two-fold formal translation, first from the *plausibility models* into the *neighbourhood models*, and the second from *neighbourhood models* into a *Kripke model*, in order for its inventory to be clearly defined within the margins of a normal logic. This is because the formal theory of belief was dynamised at this level, as opposed to

the theory of knowledge which is to be dynamised in the superstructure of the model.

(3) The superstructure, formally defining the notion of *algorithmic knowledge*, necessary for situating the discussion within the margins of *dynamic Distributed Systems Model*. The theory of knowledge was thus far captured by a static system S4.2, which served as a framework of inferable elements of the epistemic set. In other words, the logic of knowledge provides information about the limits of inferability within the epistemic context. This formal theory of knowledge maintained all the desired systematic properties, while its normality was not hindered in any way. However, within a dynamic context such as the Distributed Systems Model, we ought to define a more practical theory of knowledge, which will capture the instances of knowledge that are *de facto* computable. It will formally define an abstract notion of an algorithm, which will calibrate the system for a multi-modal, multi-agent setting.

Now that this structure is at least to some extent clear, we can venture into a more detailed analysis. When approaching the domain of modelling epistemic positions such as verificationism, one needs at all times be fully conscious of the hard-locked interrelationship between semantics and epistemology that is implied. The first question that comes to everyone’s mind is: Are we defining an epistemology through the lens of semantic meaningfulness, or are we establishing semantics on the groundwork of cognitive conditions. The answer is actually both. Our cognitive conditions for epistemic and doxastic discrimination of possible situations precisely *corresponds* to a full semantic theory when read as such.

So far we have discussed quite a few modal systems that dealt with the notions of knowledge and belief, and in this chapter I will attempt to flesh out my own take on the issue. First, I will attempt to display the necessary groundwork for formally establishing an antirealist verificationist epistemology through defining the conceptual-linguistic basis for epistemic criteria of meaningfulness, then I will show that we are able to define meaningfulness through a specific modal restriction on an accessibility relation (in particular, an indistinguishability partition structure), and I will show why it ought to be defined as a base structure of the thus-far established epistemic-doxastic model. After that, I will display the epistemic-doxastic model that I will use for modelling this specific epistemology, comprising an S4.2 logic of knowledge, and the CDL logic of belief (which serves as a dynamic equivalent of the logic KD45 which converges with the proposed logic of knowledge). Then I will attempt to overtly display the interrelationship between the epistemic-doxastic structure and the model for meaningfulness as a base structure and discuss some philosophical implications that follow from the proposed setup. After that, in order to complete the verificationist picture, I will address a necessary philosophical revision of the T axiom in form of an interpretation that should calibrate it to an antirealist epistemology. Further I will develop the formal theory further with situating the model within an abstract epistemic context, the Distributed Systems Model. The model will now be situated within a multi-agent environment in which we will be able to observe higher-order instances of knowledge and belief, along with the possible addition of operators for common and distributed knowledge. This part of the chapter will, hence, comprise the application of this combined structure of modal systems to the domain of Distributed Systems Models. We can observe the operational architecture of the DSMs as a reminder of a sort and see how they deal with the notions of knowledge, belief, and subjective indistinguishability. As DSMs are originally developed as S5 systems, I will propose the way of framing them in the weaker logics S4.2, and CDL so they model epistemic and doxastic states of agents, and will try to show how the S5 frame (which is definitionally integrated in the understanding of knowledge and belief that I propose in the model) works as a semantic screen for usable and programmable information.

All of the agents' respective behaviours, along with their epistemic and doxastic states will be explicable, and finally determined by the syntactic and semantic properties of the logics that constitute the formal structure of the model. After that, it would seem appropriate to offer some examples to see how the system behaves within the provided framework. Among the examples, we will observe an extended version of the example with Jane and the dog in the park, which will now be situated in a multi-modal and multi-agent environment.

The final part of the chapter that follows the examples is the one concerning our metalogic, i.e. the description of the system and its respective properties. I will show proofs of frame completeness, soundness, and decidability for all the logics that the model uses; S4.2, KD45 and CDL as its dynamic expansion, and S5. As not all frames were originally supported by Kripke structures (CDL system), the metalogical claims will follow mutatis mutandis from the terminology that is typical for augmented neighbourhood frames. I will again remind the reader of the restrictions on accessibility relations of each proposed system, and will elaborate in more detail what I have started in the chapter 3 – an overt explication of the interrelationship of frames through the language of set theory (following Stalnaker, 2019), as it appears to be the most natural way to think about it.

Finally, in the next chapter I will provide a short summary of the work that was done within the margins of this dissertation, along with a few remarks on how I see this project developing in potential further research. I will also attempt to argue one of the thoughts that I had in mind while writing this; when doing formal epistemology, or maybe even any kind of theoretical formal modelling, on thing must be clear – there is no single way of going about it. One formalism can support many various theories, and one theory can be supported by infinitely many formalisms. Our theories and formal models are not uniformly and unequivocally translatable one to another. The theories will always be more fine grained, and the formal structures will always enjoy a level of rigour that the theories will not be capable of attaining. Furthermore, it should always be on a theoretician's mind that the model should be philosophically or scientifically read. Only when interpretation occurs are we able to see if such and such way of defining the formal infrastructure has merit in our research.

THE CRITERION OF MEANINGFULNESS

As we have observed in the previous chapter that explicated an account of verificationist epistemology, I have attempted to offer a theoretical account which will cover the criterion of meaningfulness in a rather unorthodox manner. This pertains to a modal reading of the verificationist criterion which states that a statement is cognitively, and by extension, semantically meaningful once there exists at least one *pair* of situations, i.e. at least two possible worlds, one which forces φ and one that forces not-φ. Only when an experiment of verification is in principle constructible through cognitive conditions are we justified in claiming that the statement in question even contains a proposition. If this requirement is met, then we know that we are dealing with a meaningful statement from our natural language, and we can state that it is an element of the subset $\mathscr{L}_V$ of the set of well formed sentences $\mathscr{L}$ that expresses some state of affairs in the logical space. This criterion of meaningfulness was defined formally in the chapter VI as following:

DEFINITION (reiterated).

$$s \in \mathscr{L}_V\ (t_1) \text{ iff } \exists P \subseteq W,\ f(s) = P,\ \exists w, v \in W:$$

$$(1)\ w \sim_{t0} v,$$

$$(2)\ \forall t'\ (t_0 < t' < t_1 \Rightarrow w \sim_{t'} v),$$

$$(3)\ w \nsim_{t1} v,$$

$$(4)\ M, w, t_1 \vDash \varphi \text{ and } M, v, t_1 \vDash \neg\varphi.$$

Since we are dealing with a structure that is based on subjective indistinguishability, the logic of meaningfulness, if the reader allows me to use such nomenclature, is an S5 system, as it does not pertain to any epistemically factive content. In other words, in order to establish conditions for meaningfulness, we do not regard the present state of affairs, only the *relationship* between the statement and the worlds in which it might or might not obtain. In this respect, the model is not laden with an implicit notion of *actuality* as an indexical for some world. Instead, it only compares situations in terms of *in principle constructability* of experiments for defining the cognitive conditions for meaningfulness.

Now it might be clear that this analysis does not pertain to the classical verificationist attempts of formulating meaningfulness through observational sentences (sometimes referred to as protocol statements) that rely on the idea of theoretically unburdened statements that work as a means of establishing a theoretical language. In such traditional verificationist enterprises, all the theoretical statements that possess meaning should be reducible, or rather translatable, into statements that stem from our direct observations. This would obviously result in an epistemic theory that would render meaningless a vast amount of statements that are used in empirical sciences and even everyday discourse that obviously possess explanatory, expository, and anticipatory value in our cognitive economy.

The modal account of the formal criterion of meaningfulness that I introduced in this dissertation might be quite less chauvinistic in terms of validating such statements as meaningful. This is the case because it is not bound to direct observations, but to conceptual frameworks and their infrastructure of definability of meaningful terms. As I have previously discussed in the Chapter VI, it is directly motivated by Carnap's (1951) in-depth analysis of conceptual frameworks and meaningful use of language.

I have indicated that this would then work as a base structure of the model – a semantic screen. It can be observed as a semantic function from the natural language L that maps onto the language of meaningful statements $\mathscr{L}_V$, based on the meaningfulness criterion. Hence, every statement which fails to pass the semantic screening within the model is considered meaningless and cannot, by definition, be fed into the next layer of the structure which pertains to epistemic and doxastic states of the agents to which the system applies. In other words, each meaningless statement cannot be neither believed, nor known. I think this is quite in line with how we perceive knowledge and belief pretheoretically, even though there might appear to be counterexamples when someone is said to believe in some nonsensical statement, or rather hold a nonsensical position on some issue. This analysis would simply state that if such a person is pressed to clearly say what they believe in, they would come up short, for the simple reason that nonsensical statements are not conceptualisable. They would have to resort to a similar meaningful statement as an interpretation of the first nonsensical one, which is the point that validates the proposed criterion of meaningfulness.

EPISTEMIC AND DOXASTIC STRUCTURE OF THE MODEL

Here we arrive at the second level of the model that pertains to epistemic and doxastic states of agents to which the model pertains. This is probably the most operative part of the model, as it should clearly define what we mean when we say that we know or believe something. It should answer the question of what states are we theoretically keen on accepting as viable options, given what is known or believed by an agent. I have previously (in chapters IV, V, and VI) introduced the inventory that is relevant for this structure of the model, however, a full explication is still lacking. First we will observe the syntactic and semantic infrastructure of the model and after we will observe some metalogical properties of the systems that were used, excluding ones that were already addressed at some point within the dissertation.

To begin with, the logic of knowledge that I opted for – the system S4.2 that is adapted to the intended domain of epistemic states is modelled by an R_k accessibility relation that we have observed to be closed under the restrictions of a reflexive, transitive and strongly convergent frame. The restrictions' schematic interpretations are displayed within the chapter III of the dissertation when I introduced the logics S4.2, S5, and KD45 as a comparative case study in order to show how axioms affect, or rather determine, the accessibility relations in Kripke frames. As a reminder, the system S4.2 is a normal epistemic logic, which means that it validates the Axiom of Distribution and the inferential rule of Necessitation. The differentia specifica of S4.2 within its class of S4-based logics, as we have seen, is the Axiom .2 that states that everything that is possible to know is known to be (epistemically) possible. Formally it is presented as following:

DEFINITION (reiterated). $MK\varphi \rightarrow KM\Phi$

This axiom generates the restriction on the accessibility relation of strong convergence, as it states that for all *factively determined* possibilities, there exists one that is accessible, which is the candidate for being the actual world. We have observed in chapter III that the logic of knowledge is deductively convergent, or rather compatible with the logic KD45. Deductive convergence in this case means that all of the theorems of S4.2 appear as theorems for KD45 in the doxastic context. The logic of belief ought to be stronger than the logic of knowledge, so that belief does not collapse into knowledge because of the maximal extension of R_b over R_k. This implies that there will exist theorems of the doxastic system which will not appear as theorems of the epistemic system, but the convergence will hold in the other direction. The combined logic of knowledge and belief, determined by the axioms of S4.2 and KD45 can then be displayed with the following five theorems pertaining to both knowledge and belief (Stalnaker, 2006);

DEFINITION.

$\vdash B\Phi \rightarrow KB\Phi$	(Positive Introspection – PI)
$\vdash \neg B\Phi \rightarrow K\neg B\Phi$	(Negative Introspection – NI)
$\vdash K\Phi \rightarrow B\Phi$	(Knowledge implies Belief – KIB)
$\vdash B\Phi \rightarrow \neg B\neg\Phi$	(Consistency of Belief – CB)
$\vdash B\Phi \rightarrow BK\Phi$	(Strong Belief – SB)

These are quite familiar by now, but here they are explicated in a multi-modal context. Moreover, the theorems that generated issues for the logic of knowledge appear to be much more palatable in the combined context. This specifically pertains to the the Axiom of Negative Introspection. In the logic of knowledge it seemed to suggest that whatever is not known by an agent is known to be not known. As we pretheoretically appear to perceive belief quite differently than knowledge, this consequence appears to be more palatable. The theorem states that if I do not hold some belief, I know that I do not hold it, which seems in line with the idea that we possess a privileged insight into our belief states.

As far as the rest of the theorems are concerned, the combined theorem of positive introspection states that if an agent believes some proposition Φ, then they know that they believe it. The axiom KIB is ubiquitous in almost every combined forma theory of knowledge and belief, and simply suggests that if an agent knows something, they also believe it. The axiom of consistent belief states that if an agent believes some proposition Φ, then they disbelieve its negation. It can also be understood that the proof of decidability relies on the system validating this theorem. Finally, one of the most important theorems of this combined enterprise gives us a philosophically interesting insight into the relationship of knowledge and belief. This is the theorem of strong belief, which states that if an agent believes that Φ, they also believe that they know it. It suggests that justification plays a role in endorsing beliefs, as we form them with the assumption of their probable factivity. Philosophically speaking, it states that we would not have formed a belief if we did not operate under the supposition of it being true. This is also the upshot of endorsing the axiom of factivity in our pure logic of knowledge, which can be understood as a representation of the philosophical position of epistemic infallibilism.

Along with the stated theorems, there appears to be quite an appropriate bridge axiom, or rather an interaction axiom, that stems from this combined formal theory of knowledge and belief. It states that belief is logically equivalent to possibility of knowledge. It formal iteration is defined as following:

DEFINITION. $\vdash B\Phi \leftrightarrow MK\Phi$

In this theorem M is a derived operator of epistemic possibility, defined through the knowledge operator as $\neg K \neg$. Hence, we can read the statement as 'iff an agent were to believe some proposition Φ, then they do not know that the do not know it'. A more natural reading would be in line with the original formulation, which states that believing some proposition Φ is equivalent to Φ being a candidate for being known. It seems that this formal consequence converges with our pretheoretic conception of the interaction between knowledge and belief, as we want our beliefs to be such that they have the potential of constituting knowledge.

Now that we have observed the inventory of this combined model of knowledge and belief, we can formally represent it as following:

$$M = \{W, R_k, R_b, [\![\ ]\!]\}$$

The set of possible worlds W and the valuation function $[\![\ ]\!]$ are standard in their formal interpretation, while the relations of accessibility are defined separately as R_k for the set of worlds accessible with respect to what is known, and R_b as the set of worlds compatible with what is believed by the agents of the system.

As I have indicated throughout this dissertation, the plan for this endeavour was to dynamise both knowledge and belief in a manner that they are calibrated to describe, or rather model, the behaviour of agents in a multi-modal, multi-agent settings which would simulate more closely epistemic and doxastic situations for resource-bound agents without the capacity to store infinite information and process them instantaneously. The dynamisation of both logics of knowledge and belief would then relieve the system of the consequence of logical omniscience, which is still present in the static model that is validated by an S4.2 and KD45 system that we have observed thus far in the chapter.

Albeit the dynamisation of both knowledge and belief is necessary in order to situate them within a verificationist reading of the framework of the Distributed Systems Models, we might want to go about it separately. I have opted to dynamise the logic of belief with a modal system that relies on conditional structures as system's atoms that we have observed in the chapter V on belief revision. Conversely, epistemic states will be dynamised through the notion of algorithmic knowledge that relies on the fact that not all knowledge that is in principle inferable is accessible to the agents within the system. This divergence in dynamising knowledge and belief separately was purely theoretically and formally motivated and bears no practical effect on the behaviour of the agents. It was made because the separate framing of the two notions appears to have the most natural reading respectively, while maintaining the desired formal metalogical properties.

DYNAMISATION OF BELIEF

As I have already indicated, the system of Conditional Doxastic Logic is not itself dynamic, but it has a dynamic reading in the philosophical interpretation. This will allow the system to maintain the desired properties, while still doing the necessary work for explicating changes in doxastic states of the agents. As seen in the fifth chapter, it relies on the fact that belief are adopted on the basis of some body of given information. Moreover, it is important to reiterate that the CDL system provides additional structure in relation to its static counterpart KD45, as it allows for defining a plausibility ordering. This calibrates agents to assess situations in terms of their likelihood, making them adaptable to their perceived environments. Along with that, the agents' states are more fine-grainedly programmable for action, as not all doxastically viable options play the role of equal importance in their cognitive economy. This is also quite in line with the verificationist theory that this system is set out to support, as the agents do not respond to the state of affairs in the world, but to their interpretation of the input data. Some agents might not be sensitive to specific variables to which other participants in the system are, so they will behave appropriately to set of interpreted information that they gather from the perceptive apparatus.

In its basis, the system CDL does not differ axiomatically from the system KD45 at all, as it validates the same postulates, albeit defined through necessarily conditional syntax. Then, if we were to dynamise the belief states of agents by using a full preordering of CDL, then the accessibility relation will be thusly defined. Nota bene, before we adapt CDL's accessibility relation to Kripke structures, it will be naturally defined in the terminology of neighbourhood structures;

DEFINITION. $M_{CED} = \{ W, R_k, N_{CDL}, [\![\,]\!] \}$

Thus, the R_k relation remains the same, as the logic of knowledge remains S4.2, while the R_b relation is exchanged for a N_{CDL} membership relation, supported by a neighbourhood structure. Now, as we have seen in the chapter V on belief revision, we know that the logic CDL validates both the Distribution Axiom K, and the rule of Necessitation, rendering the neighbourhood frame augmented. As there is a proof in (Negri and Pavlović, 2023) paper on proof theoretic reading of CDL, all augmented neighbourhood structures are translatable into Kripke frames, making the model validated by a set of two normal systems, one for knowledge and the other for belief. The updated model, once we have secured its normality, can be displayed as the following Kripke structure;

$$M_{CED} = \{\, W, \mathrm{R_k}, \mathrm{R_{\leq CDL}}, [\![\;]\!] \,\}$$

Finally, before we venture into the discussion on algorithmic knowledge and adaptation of the logical model to the framework of Distributed Systems Models, one thing remains on the table. In order for our model to be completely adapted to the verificationist enterprise, we ought to adapt the reading of the axiom of factivity within the logic for knowledge S4.2. The axiom of factivity is formally defined as following, and provides a restriction of the accessibility relation of reflexivity on the frame;

DEFINITION. $K\Phi \rightarrow \Phi$

The axiom itself states a simple relation between knowledge and truth, which is completely unproblematic within the margins of a realist reading. However, if we want to talk about the notion of truth in a verificationist setting, we appear to be in need of a modifier. The classical reading of the T axiom states that everything that is known is also true, but in order for this framework to capture the essence of verificationism, we ought to read it as 'everything that is known is *demonstrably true*. This does not in any way constitute a formal revision of the logical apparatus that the model validates, only its interpretation for the sake of a more natural extrasystematic reading.

ALGORITHMS FOR KNOWLEDGE AND ACTION

At this point in the dissertation we have observed an explication of the second layer of the model used for capturing a verificationist epistemology in a multi-modal setting, but in order to complete the model with the dynamisation of knowledge through algorithmic devices, we ought to adapt the terminology to fit the new framework of Distributed Systems Models. This will constitute the final rendition of the model, which will become adapted to non-omniscient knowers in a multi-agent environment.

As we have observed in the chapters III and IV on epistemic and doxastic logics, Distributed Systems Models, and algorithmic knowledge, when situating epistemic and doxastic modelling in the context of Distributed Systems Models, we will adapt the relevant terminology. The agents are understood as interconnected processors with finite numbers of mutually exclusive and jointly exhaustive states. There is a non-nought number of NULL processors that provide input about the environment to the processors of the system. Since we are in an anti-realist setting, the NULL processors do not display the state of affairs in the world, but a set of codified pieces of data, i.e. the way the processors perceive their surroundings. The processors being fed only cognitively operative information is at the very hear of both the verificationist understanding of meaning and the formal notion of intensionality. The intensional aspect of this framing is the coarse-grained discriminative aptitude of processors to classify information in worlds that are different from one another in some relevant aspect with and S5 validating relation of accessibility. In other words, if the NULL processors provide them with inputs from which they are incapable of discriminating in which world they are, the input itself does not play a role in their cognitive economy. This is simply another way to define the meaningfulness relation explicated with the base structure of the model, with adapted terminology of Distributed Systems Models.

Furthermore, as their logics of knowledge and belief are S4.2, and CDL, respectively, the sets of worlds they discriminate over are aligned with the accessibility relations supported by the systems in question. This means that the worlds that are epistemically accessible do not collapse into worlds that are doxastically accessible and vice versa, as the structures of accessibility for knowledge and belief are distinctly defined. The partial-preordering of S4.2 defines the set of global states that are epistemically accessible, while the full pre-order with a plausibility relation of the CDL defines the set of global states that are doxastically accessible.

As the Distributed Systems are structurally dynamic models, we infuse them with an algorithmic function for computable knowledge displayed in the chapter IV of the dissertation. As the name suggests, algorithmic knowledge is defined as de facto computable knowledge that the processors are capable of inferring, given the corpus of information within the local state they are in, and the algorithmic procedure for computing the proposition in question. This constitutes the third layer of the model, or the *superstructure*, which is built on the epistemic-doxastic structure. As we have dynamised the doxastic partition of the model within the second layer of the model, the superstructure pertains solely to the epistemic partition. We have also observed how the system behaves when runs and time are added to the structure, as we are able to track changes in global and local states of the processors with respect to dynamic behaviour of the system. The formal reconstruction of this notion was defined as following:

DEFINITION (reiterated).

$$(I, r, t) \vDash X_i\varphi \text{ iff } A\langle alg_i, c\rangle\,(r, t) = \text{“Yes”}, \text{ for } alg_i\,(r, t), \text{ and } c = data_i\,(r, t)$$

The interpreted system I in the context of the run r, at time t models algorithmic knowledge of the proposition φ – X_iφ iff there exists a local state, comprising an algorithm alg_i and the informational state c for the same run and time of the system, that answers "Yes" to the question of whether φ obtains. Note that the question-answer basis for the definition has little to do with linguistic competence, but is a mere idealisation of how the processors would represent their knowledge if they were in a linguistic setting. Their knowledge and beliefs are, as was the case until now, manifested through their actions and interactions.

With the added function of algorithmic knowledge, the model is now calibrated to differentiate between everything that is in principle inferable within the system and everything that each processor is capable of inferring within the system in a certain run r and time t. Furthermore, this dynamisation allows us to theoretically easier differentiate between processors, as not all of them possess the same local states that are defined with algorithms and contextual information. With this formal instruments, we are able to capture processor's mono-agent multi-modal epistemic and doxastic states of any order, along with their criterion of meaningfulness, established on the base structure of the logical model. The final step for the model to be fully immersed and adapted to the Distributed Systems verificationist setting, we ought to find a way to model two types of multi-agent knowledge. It also ought to be stated that this final addition is supported by the same structures, with the axiomatic schemata and theorems being minimally adapted *mutatis mutandis*.

The first type is *common knowledge*, represented as C_Gφ, expressing the type of knowledge usually found in PAL systems and often displayed and discussed with muddy children-like examples (Chapter IV). Every child is given the same piece of information, but not only that, each of them know that everyone got the same new piece of information. So, public announcement of facts results in common knowledge of agents in the system. As we are not in the classical DEL formal environment, I will have to adapt the definition to fit the framework we are observing. The second type of multi-agent knowledge, which is even more philosophically and formally interesting within such contexts is *distributed knowledge*, represented by D_Gφ. This type of knowledge pertains to situations in which none of the agents are independently apt to infer some information φ from the given informational context. This is not necessarily caused by their lack of computational capacity, but they *individually* might lack sufficient data for the inference in question. However, once their collective knowledge is put to use, along with their collective algorithms, they are capable of inferring φ.

As the local states of agents in DSMs are now finer-grained than in the original formulation, we are able to formally capture common and distributed knowledge by using constituents of local states of the agents, i.e. processors. Thus, as a final expansion to the model, we add the following definitions:

DEFINITION.

$M, w \models C_G\varphi$ iff $\exists g_n$ such that $(\cap\{A_0\langle alg_0, c_0\rangle, A\langle alg1, c_1\rangle, \langle alg2, c_2\rangle, \ldots \langle alg_n, c_3\rangle\} \models \varphi)$

Common knowledge is here defined as following; There exists common knowledge of the proposition φ in an epistemic situation if the epistemic situation is such that the intersection of local states of the agents of the system is capable of modelling φ. As we are using the intersection operator, all of the processors ought to be capable of modelling φ with respect to their given contextual knowledge and algorithmic competence. As opposed to common knowledge, distributed knowledge is represented in the following fashion:

DEFINITION.

$M, w \models D_G\varphi$ iff $\exists g_n$ such that $(\cup\{A_0\langle alg_0, c_0\rangle, A\langle alg1, c_1\rangle, \langle alg2, c_2\rangle, \ldots \langle alg_n, c_3\rangle\} \models \varphi)$

Instead of using the intersection operator to define this type of knowledge, I propose using the union operator, as for distributed knowledge, there ought to be sufficient *combined* background knowledge and algorithmic competence of the group in order to model φ in a certain epistemic situation. We can intuitively read the formal explication as follows; A group has distributed knowledge that φ if and only if there exists an epistemic situation in which the union of local states (contextual information and algorithms) of the agents is sufficient for modelling φ.

Both of these multi-agent types of knowledge are at the very heart of the Distributed Systems and verificationist epistemology, as they enrich the model with one of the most important features for establishing a formal epistemic and doxastic theory – *information location* within the system. Furthermore, the system is now calibrated to deal with processors' interactions through time. Their common and distributed knowledge will change over time, as they are communicating and learning. This rendition of the model is far removed from our initial static conception of DSMs, as the systems' participants are immersed in a temporal, inferential context, and their knowledge and belief can be modelled in a less idealised fashion. Finally, it ought to be stated that the intended domain of DSMs application has not changed at all – if we were to observe a highly idealised non-dynamic epistemic setting, this model would be just as applicable. We would simply ignore the variables that are not operative within the model.

In the next chapter, I will briefly talk about some metalogical properties of the systems that we have been using. The focus will be on the logic of knowledge, S4.2, as the doxastic logic CDL was already elaborated on in detail in the fifth chapter on belief revision and epistemic update.

CHAPTER VIII – METALOGIC

LOGIC S4.2 AND ITS PROPERTIES

Now it seems that it would be appropriate to designate this part of the dissertation to the epistemic logic that I opted for in the dissertation – the logic S4.2. To be more precise, it is the logic that underpins the R-accessibility relation that I use within the model in order to formally define knowledge operator and its entailment relation, along with the Rs relation that I use for subjective indistinguishability and the weak information ordering relation for beliefs (supported by CDL axioms in a neighbourhood frame).

As I hope was clear from the title of this dissertation, along with the discussion of the case studies of logics in the Chapter 3, S4.2 is a normal modal system in the range of logics between S4 and S5 that validates (1) all the propositional variables of the Classical Propositional Logic, (2) the K schema, also known as the Distribution Axiom, (3) the rule of Necessitation, (4) the rule of uniform substitution (if needed, with respect to framing), (5) the Axiom 4, i.e. the Axiom of Positive Introspection, with (6) the Axiom .2 as its differentia specifica in the context of its class of systems. The system S4.2 itself was first put on the scene by Michael Dummet and E.J. Lemmon while they were working on the modal systems within the range of S4 to S5. (Chalki et al., 2018, p.1) The system was since advocated as the correct logic of knowledge by theoreticians such as Lenzen (1978) and Stalnaker (2006), but I see no point in proclaiming an epistemic model (or the logic that it validates) as correct outside the meaning of its adequate application.

Outside of the intended domain of epistemic and doxastic logics, the system S4.2 was shown to be useful for modelling relativistic spacetime when applied to the inteded domain of temporality, along with establishing the basis for Einstein's theory of special relativity in the context of Minkowskian four-dimensional geometry. (Chalki, et al. 2018) Another common name for this system, as the authors discuss, is the logic KT4G, as the .2 axiom is also called the G axiom for some authors. (Chellas, 1983; Takano, 2019) Seeing as the logic is normal, it is supported by the structure of Kripke frames, which necessitates no full orderering relations, but solely a linear distribution of worlds through a partial pre-ordering, as is the case for the systems belonging to its class. To qualify what that means, we might want to take a brief look into the accepted terminology of classifying modal frames. A partial ordering is a frame that is closed under reflexivity, transitivity, and anti-symmetry. A partial pre-ordering is a frame that is closed under reflexivity and transitivity, while a strict order is closed under irreflexivity and transitivity. (Chalki, et al., 2018)

In order to refresh the readers' memory, when read in the intended domain of epistemic logic, the defining Ax.2 or the G axiom for the logic in question basically states that everything that is possible to know is also known to be possible. Formally, we can represent it as following;

DEFINITION Ax.2 $MK\varphi \rightarrow KM\varphi$

The accessibility relation of its class of systems (S4 – S5) are all closed under reflexivity and transitivity (and some weaker properties that follow by extension, such as seriality, and shift-reflexivity), however, the newly added restriction on R that is added to ones that were already implied is convergence. This condition can also be found in the relevant literature under the name of weak directedness (Chalki et al., 2018, p.2). The condition of convergence can also be described as weak and strong, and both can be proved to follow from the proposed .2 axiom (Stalnaker, 2016), but that distinction does not pertain this paper, as it is only relevant when the model in question uses first-order quantification. For our purposes the distinction plays no role, as the model devised in the dissertation uses coarse-grained semantics, and hence, propositional atoms as the basis of our language.

As opposed to the S5 system, the S4.2 system possesses many coveted properties such as interdefinability of knowledge and belief. When the logic of choice for knowledge, while constructing a model, is the S5 logic, then belief collapses into knowledge, as seen in the chapter 3 of this dissertation. As seen in Lenzen's and Stalnaker analysis, when using the system S4.2 for knowledge, and employing an additional bridge axiom for belief that states

DEFINITION. $B\varphi \leftrightarrow \neg K \neg K\varphi$,

we acquire a congruent hybrid model of knowledge and belief, underpinned by two logics, S4.2 for knowledge and KD45 for belief with all appropriate properties. As we have seen, the accessibility relation that is entailed by the logic of knowledge within the epistemic model must not be defined through a proper subset of the logic of belief, because that would imply that we only have true beliefs about the world, and furthermore, that we might not have accessible some beliefs about some true facts in the world. The first consequence is obviously preposterous, while the second is at least problematic for any antirealist epistemology. Following Stalnaker's interpretation, we read the hybrid axiom as presented: "One believes that φ, in the strong sense, if and only if it is compatible with one's knowledge that one knows that φ."

The detailed analysis of defining such a hybrid model that Chalki et al. (2018) discuss when revisiting Lenzen's and Stalnaker's methodology can be summarised as following:

(1) If we were keen on accepting the axiom that states that knowledge implies belief

DEFINITION. $K\varphi \rightarrow B\varphi$ (KIB)

(2) Then the axiom that we can call (for the lack of a universally accepted name) the axiom of negative strong belief

DEFINITION. $B\varphi \rightarrow \neg B \neg K\varphi$ (NSB)

That virtually states that we do not believe anything that we don't believe we know,

(3) And the axiom of positive introspection of having beliefs

DEFINITION. $B\varphi \rightarrow KB\varphi$ (PIB)

Combined with the bridge axiom that interdefines knowledge and belief that I stated earlier, we acquire the S4.2 system that pertains to the epistemic part of the model, and KD45 system that pertains to the doxastic part. Following Lenzen (1978) with adapted terminology for intertextual consistency reasons, we can state formally

DEFINITION. KD45 + S4K + KIB + NSB + PIB = S4.2KB

Again, the KD45 logic will impose constraints on accessibility relation for the B operator (hence defining what belief means for the system), and S4.2 logic will do the same for knowledge, mutatis mutandis. As both Stalnaker (2019) and Chalki (2018) notice, this approach naturally captures Williamson's (2000) knowledge first programme, as it appears that we have derived the meaning of belief via the meaning of knowledge, but I will not delve deeper into this observation, as it appears to me that the model could just be constructed upside-down starting from belief as the contents of it are purely definitional. Chalki et al. (2018), however, notices that belief wouldn't be definable through knowledge if we left out the axiom NSB.

Stalnaker has replicated Lenzen's results in his monograph Knowledge and Conditionals (2019), with added philosophical implications that he felt provided more reasons for accepting S4.2 logic as a good candidate for our static logic of knowledge.

As far as metalogical properties, as seen in Tiomkin and Kaminski (1996), S4.2 is a system that is strongly complete by belonging to a class of logics with a final non-empty cluster – the consequence of validating strong convergence restriction on the Rk accessibility relation. Let us start with their proof of soundness;

PROOF. Let us consider F, an arbitrary frame that is closed under reflexivity and transitivity such that its final cluster is non-empty. The axioms that validate reflexivity and transitivity, viz. positive introspection for K obtain by definition. As for the axiom .2, we can take any model M established on the arbitrary frame F and an arbitrary w element of W such that it validates ◇□P. It follows that there must exist a world vEV that is accessible from w and it validates □P.

Let u be an arbitrary world accessible from w (wRu). Since we presupposed a final non-empty cluster, let us consider an arbitrary world s that is itself the element of the final cluster. As it follows that wRu and uRs, we are justified to conclude from the fact that M, s satisfies p, that M, u satisfies ◇P, so as u was introduced as an arbitrary world that was accessible from w (wRu) it follows that M, w models □◇P.

CDL LOGIC

CDL, as opposed to S4.2 was formally fleshed out in the chapter VI of the dissertation, so we can briefly summarise what has been so far observed. CDL is a dynamic variation of the KD45 logic that adopts conditional structures as its atoms. I have thus far stated that it is not a classically dynamic logic, as there are no standard dynamic PAL-type operators defined within its language, however, it is able to express dynamic attitudes of agents within the system in an indirect way; we can view the conditional structure as a function. The function basically attributes a proposition to a corpus of knowledge (or belief) by defining the doxastic situation in which it is acceptable. The conditional structure (A|B) can thus be read as A being acceptable once B is accepted. This is obviously sufficient to interpret the system dynamically, notwithstanding the static basis of the system.

We have observed that CDL was originally explicated as a mono-agent system, however, this can be remedied by trivially calibrating the language and it atomic structure to a non-empty group of agents that bare conditional doxastic states. Furthermore, such an expansion does not affect in any way the properties of the system. (Board, 2004). The detailed and extensive proofs of soundness and completeness can be found in (Baltag, et al. 2015). The said proofs cover both versions of CDL, the mono-agent and multi-agent.

Apart from its axiomatic equivalence to KD45, it is also of importance to revisit the fact that CDL works as a full formal explication of AGM-style theories, with concrete syntactic infrastructure. This is exactly what makes it a good candidate for pairing it with an algorithmically dynamised S4.2 system, as the belief revision it provides is not in theoretical or formal tension with the epistemic update that the algorithmic variation of S4.2 provides.

Finally, in order to remind the reader, I have stated that the framing of CDL was initially defined within a plausibility model structure, that is explicable within the context of a neighbourhood structure, in which we can describe the instruments of CDL such as plausibility ordering through the notion of nested spheres in the terminology of membership relations, typical for neighbourhood structures. Moreover, CDL validates both the Distribution axiom and the rule of Necessitation, which guarantees that its structure is to be understood as an augmented neighbourhood. Consequently, an augmented neighbourhood can be supported by a Kripke frame, rendering the doxastic aspect of the combined epistemic-doxastic model normal.

CHAPTER IX – CONCLUDING REMARKS

In this dissertation I have attempted to show how a formal model can be devised in an endeavour of fleshing out and supporting an informal epistemological theory. The formal model in question uses the modal notion of possible worlds in order to capture what epistemic and doxastic situations can be understood as viable options for an agent at a certain point in time. The motivation for such an approach is devising methods that provide us with sufficient nuance and clarity on one hand, while maintaining theoretical robustness on the other. When applying epistemic and doxastic models to non-ideal agents, we are able to detect the problematic aspects of an epistemology or an underlying logic, both of which often generate unpalatable consequences that we might not be aware of before considering them in an abstract formal setting. The model that I have built here was designed to support a verificationist theory of knowledge, belief, and meaning, and so its elements were calibrated to the terminology of such theoretical accounts.

The second chapter of the dissertation deals with our everyday use of modalities and modal statements in which I have introduced some peculiar examples that might have been proven to be useful in the more technical parts of the book. We have briefly observed the role that conditional structures have in modal reasoning, taken discussed some of them in terms of their theoretical underpinning.

The third chapter covered the introductory part to epistemic and doxastic logics that were meant to support epistemic and doxastic terminology from the informal theoretical background. I have opted to take a case study of three discrete systems, viz., the S4.2 system in its epistemic interpretation, the S5 system with the same intended domain, and finally the KD45 system in its doxastic interpretation. I have analysed all of the said systems via the operational part of their syntactic infrastructure, the accessibility relation, presented in and supported by a Kripke frame. Furthermore, I have provided a schematic interpretation for the systems as a visualisation that might be proven to be helpful in order to get a better grasp on the linearity of the relations. All of the logics that were analysed so far have been displayed in their static variants. The final part of the chapter introduces a type of an externalist epistemic model, usually entitled Distributed Systems Model. It works as a sort of a framework for dealing with complex, uncentralised structures with a set of interconnected agents within them that interact in ways bound by logical and extra-logical rules. Their epistemic and doxastic states are determined by an abstract theoretical notion of a 'local state' that can be understood dispositionally in terms of their engagements with their environments.

The first dynamisation that we observe within the margins of this dissertation takes place in the fourth chapter, in which I introduce an algorithmic notion of knowledge that establishes a differentia specifica between what is in principle inferable within the system and what the agents are actually capable of computing or inferring, given their background information and the procedures for inference that are at their disposal in certain epistemic and doxastic situations. Within this framework, algorithms are understood as agents' procedural capacities to problem-solve in various abstract situations. Theoretically they work as a sort of a buffer for hindering the problem of logical omniscience, which is here defined as closure of epistemic and doxastic operators under the material implication or logical equivalence. This is the main motivation behind introducing dynamics to the model, although as we observed in the following chapter, they are not our only resort when dynamising a system. Furthermore, when dealing with a system that recognises algorithmic knowledge, we are capable of inspecting and understanding the local states of the agents in more detail; they become definable as a pair of (1) background information that the agent possesses in a certain epistemic situation and (2) an inferential procedure they can take in order to explicitly arrive at some information that can enable them to communicate with their environments more successfully.

In the fifth chapter of the dissertation I introduce a variant of the doxastic logic KD45 that was observed in the third chapter when discussing the case studies of logical systems and their properties. This variant, the system CDL, is based on conditional atoms, while retaining all of the structural properties and the axiomatisation of the static logic KD45. It is a logic that is based on a linear plausibility ordering that can help differentiate between soft and hard informational inputs. In other words, it establishes a notion of epistemically indistinguishable worlds, which can be ordered by plausibility when the agents are presented with non-conclusive evidence for some fact within the system. The original plausibility ordering model is then translated into a neighbourhood model that can define the ordering with the terminology of Grove's nested spheres. Once it is established that the proposed system of CDL validates both the axiomatic scheme of Distribution and the rule of Necessitation, the model is translated to a Kripke frame, guaranteeing its normality.

The sixth chapter deals with the informal epistemic theory that the model was set out to support a variant of verificationism, whose criterion of meaningfulness is reframed to a modal approach. The criterion is, hence, defined as a sort of a cognitive test of statements in a manner that an agent is capable of conceptualising procedures for determining the existence of a possible world in which the statement is true, and one in which it is false. Any and all statements that do not meet the criterion are deemed meaningless and hence, play no role in the cognitive economy of the agent in question. The notion of meaningfulness is also presented akin to scientific exploration, as a statement can be deemed meaningless by an agent at a point t_0, but meaningful at another point t_1, as the methods for capturing variables in the agents' environment can develop diachronically. This theoretical account can be viewed as an extension of Carnap's ideas proposed in the article Empiricism, Semantics, and Ontology (1951.).

The following chapter can be understood as a synthesis of every aspect of formal epistemic and doxastic modelling we have thus far observed. Within it I proposed a complex model, comprising three structural layers, (1) the base layer of semantic and cognitive meaningfulness, (2) then the layer of epistemic and doxastic notions, working as a basis for defining their behaviour and interrelationship, and (3) finally, the layer of algorithmic knowledge, reconstructing the type of information that works as an input for the system, and the way the agents are capable of manipulating and reframing it in order to be better in their communication with their environments.

The final chapter of the dissertation is dedicated to the metalogical properties of the systems that the epistemic and doxastic model uses, viz. S4.2 and CDL for knowledge and belief, respectively.

The field of formal epistemology is developing at an impressive speed, as some topics such as dynamic and agentive modalities have gained traction in the last decade. The potential for further research, at least in my view, exists especially in the development of multi-modal, multi-agent, dynamically oriented systems that are capable of defining the intricacies that we encounter when analysing, observing, and examining the notions of knowledge, belief, and meaningfulness. Work on formally defining motivational states through reason-like states of agents within the DSMs could provide a new basis for understanding knowledge accessibility, and even more importantly, knowledge usability. A game theoretic approach to STIT-like systems can be used for developing machine learning and calibrating the notion of rationality to various standards.

REFERENCES

Bennett, J., 2003. *A Philosophical Guide to Conditionals*. Oxford: Oxford University Press.

Berto, F., French, R. & Priest, G., 2018. *Modes of Truth: The Unified Logic of Truth and Modality*. Oxford: Oxford University Press.

Berto, F. & Jago, M., 2019. *Impossible Worlds*. Oxford: Oxford University Press.

Biggs, S., 2011. Ifs without cans. Philosophical Studies, 154(3), pp.383–400.

Bordini, R.H., Dastani, M., Dix, J. & Seghrouchni, A.E.F., eds., 2006. Multi-Agent Programming: Languages, Platforms and Applications. New York: Springer.

Carnap, R., 1951. Empiricism, semantics, and ontology. Revue Internationale de Philosophie, 4(11), pp.20–40.

Carnap, R. 1932, The elimination of metaphysics through logical analysis of language, *Erkenntnis*, vol. 2, pp. 219–241.

Chalki, A., Koutras, C.D. and Zikos, Y., 2018. A quick guided tour to the modal logic s 4. 2. *Logic Journal of the IGPL*, *26*(4), pp.429-451.

Chellas, B.F., 1980. *Modal Logic: An Introduction*. Cambridge: Cambridge University Press.

Doyle, J., 1979. A truth maintenance system. Artificial Intelligence, 12(3), pp.231–272.

Dummett, M. 1978, *Truth and Other Enigmas*, Duckworth, London.

Edgington, D., 1995. On conditionals. Mind, 104(414), pp.235–329.

Einstein, A. & Rosen, N. 1937, 'On gravitational waves', *Journal of the Franklin Institute*, vol. 223, no. 1, pp. 43–54.

Fagin, R., Halpern, J.Y., Moses, Y. & Vardi, M.Y., 1995. *Reasoning about Knowledge*. Cambridge, MA: MIT Press.

Fine, K., 2015. Angellic content. Journal of Philosophical Logic, 44(6), pp.613–649.

Fitch, F.B. (1952) 'A logical analysis of some value concepts', *Journal of Symbolic Logic*, 17(2), pp. 105–118.

Frege, G. (1953) *The Foundations of Arithmetic: a logico-mathematical enquiry into the concept of number*. Translated by J.L. Austin. 2nd rev. edn. Oxford: Basil Blackwell.

Garson, J.W., 2006. *Modal logic for philosophers*. Cambridge: Cambridge University Press.

Gettier, E.L., 1963. Is justified true belief knowledge? Analysis, 23(6), pp.121–123.

Haas, G., 2015. *Minimal Verificationism: On the Limits of Knowledge*. Berlin, München, Boston: De Gruyter.

Halpern, J.Y., 2007. *Reasoning about Uncertainty*. Rev. ed. Cambridge, MA: MIT Press.

Halpern, J.Y., Moses, Y. & Vardi, M.Y., 1994. Algorithmic knowledge. In R. Fagin, ed. Theoretical Aspects of Reasoning about Knowledge: Proceedings of the Fifth Conference (TARK 1994). San Francisco: Morgan Kaufmann, pp. 255–266.

Heraclitus, 2001. Heraclitus: Fragments. Translated by T.M. Robinson. Toronto: University of Toronto Press.

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

Horty, J.F., 2000. *Agency and Deontic Logic*. Oxford: Oxford University Press.

Jago, M., 2014. *The Impossible: An Essay on Hyperintensionality*. Oxford: Oxford University Press.

Kaminski, M. & Tiomkin, E., 1996. The expressive power of second-order propositional modal logic whose modalities are S4.2 or weaker. Notre Dame Journal of Formal Logic, 37(1), pp.64–86.

Kment, B., 2012. *Modality and Explanatory Reasoning*. Oxford: Oxford University Press.

Kratzer, A., 1991. Modality. In A. von Stechow & D. Wunderlich, eds. Semantics: An International Handbook of Contemporary Research. Berlin: de Gruyter, pp.639–650.

Kratzer, A., 2012. Modals and Conditionals: New and Revised Perspectives. Oxford: Oxford University Press.

Lenzen, W., 1978. Recent Work in Epistemic Logic. Dordrecht: Reidel.

Lewis, D., 1963. *An analysis of knowledge and valuation*. Ph.D. thesis. Harvard University.

Lewis, D., 1996. Elusive knowledge. Australasian Journal of Philosophy, 74(4), pp.549–567.

Lewis, C.I., 1918. *A survey of symbolic logic*. University of California press.

Misak, C.J. (1995) 'The logical positivists and the verifiability principle', in Misak, C.J. (ed.) *Verificationism: Its history and prospects*. New York: Routledge, pp. 1–28.

Nagel, T. 1986, *The View from Nowhere*, Oxford University Press, New York.

Negri, S. and Pavlović, E., 2024. A proof-theoretic approach to formal epistemology. In *Saul Kripke on Modal Logic* (pp. 303-345). Cham: Springer International Publishing.

Pacuit, E., 2017. *Neighborhood Semantics for Modal Logic*. Cham: Springer.

Parmenides, 1983. Parmenides: On Nature. In G.W. Most, ed. Early Greek Philosophy, Volume V: Western Greek Thinkers, Part 2. Cambridge, MA: Harvard University Press, pp. 111–183.

Pettigrew, R. & Weisberg, J., eds., 2019. *The Open Handbook of Formal Epistemology*. PhilArchive.

Popper, K.R., 1959. *The Logic of Scientific Discovery*. London: Hutchinson.

Priest, G., 2017. *Towards Non-Being: The Logic and Metaphysics of Intentionality. 2nd ed.* Oxford: Oxford University Press.

Stalnaker, R.C., 1968. A theory of conditionals. In N. Rescher, ed. Studies in Logical Theory. Oxford: Blackwell, pp.98–112

Stalnaker, R., 1999. *Context and Content: Essays on Intentionality in Speech and Thought*. Oxford: Oxford University Press.

Stalnaker, R., 2006. On logics of knowledge and belief. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, *128*(1), pp.169-199.

Stalnaker, R., 2019. *Knowledge and Conditionals: Essays in Honor of Ernest Sosa*. Oxford: Oxford University Press.

Takano, M., 2019. A modified subformula property for the modal logic S4. 2. *Bulletin of the Section of Logic*, *48*(1), pp.19-28.

van Benthem, J., 2007. Dynamic logic for belief revision. Journal of Applied Non-Classical Logics, 17(2), pp.129–155.

van Fraassen, B.C., 1980. *The Scientific Image*. Oxford: Clarendon Press.

Veltman, F., 1996. Defaults in update semantics. Journal of Philosophical Logic, 25(3), pp.221–261.

Williamson, T., 2000. *Knowledge and Its Limits*. Oxford: Oxford University Press.

Wittgenstein, L. (1961) *Tractatus Logico-Philosophicus*. Translated by D.F. Pears and B.F. McGuinness. London: Routledge & Kegan Paul.

Wooldridge, M., 2009. *An Introduction to MultiAgent Systems. 2nd ed*. Chichester: Wiley.

Xenophanes, 1994. Xenophanes of Colophon: Fragments. Translated by J.H. Lesher. Toronto: University of Toronto Press.

SCHEMATIC INTERPRETANTIONS (ILLUSTRATIONS) LIST

ILLUSTRATION.1

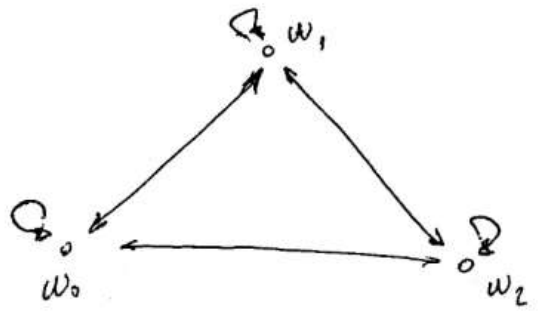


ILLUSTRATION.2

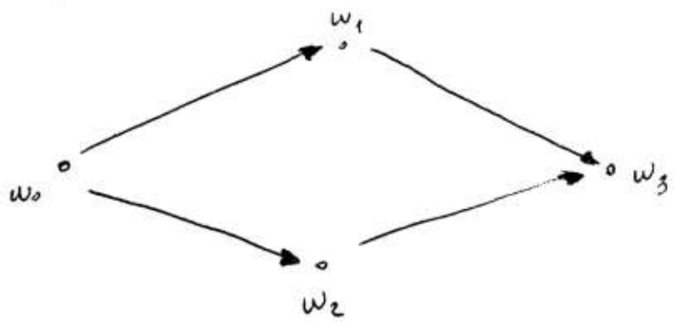


ILLUSTRATION.3

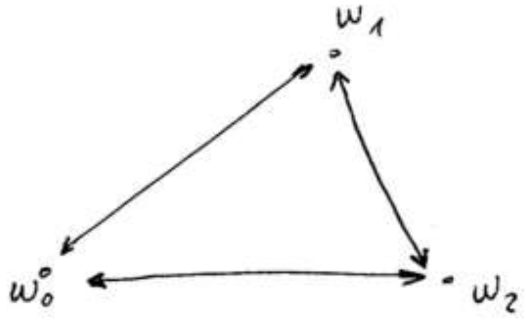


ILLUSTRATION.4