Epistemic Modality, Mind, and Mathematics

Hasen Khudairi

June 20, 2017

1

Abstract

This book concerns the foundations of epistemic modality. I examine the nature of epistemic modality, when the modal operator is interpreted as concerning both apriority and conceivability, as well as states of knowledge and belief. The book demonstrates how epistemic modality relates to the computational theory of mind; metaphysical modality; deontic modality; the types of mathematical modality; to the epistemic status of undecidable propositions and abstraction principles in the philosophy of mathematics; to the apriori-aposteriori distinction; to the modal profile of rational propositional intuition; and to the types of intention, when the latter is interpreted as a modal mental state. Examining the nature of epistemic logic itself, I develop a novel approach to conditions of self-knowledge in the setting of the modal μ-calculus, as well as novel epistemicist solutions to Curry's and the liar paradoxes. Solutions to the Julius Caesar Problem, and to previously intransigent issues concerning the first-person concept, the distinction between fundamental and derivative truths, and the unity of intention and its role in decision theory, are developed along the way.

Acknowledgements

othy Williamson, Jessica Wilson, Keith Wilson, Crispin Wright, Stephen Yablo, and Seth Yalcin.

From 2005 to 2008, I was an honors undergraduate in philosophy at Johns Hopkins University. For their encouragement and example, I am grateful to Michael Williams, Hent de Vries, Meredith Williams, Dean Moyar, and Maura Tumulty.

For his friendship and for visits at the beginning and end of the time this dissertation was written, I am grateful to Armand Leblois. For their unwavering support of my endeavors over the years, I am grateful to my parents.

The book is dedicated to Alison Bowen.

Chapter **5** is forthcoming as "Conceivability and Haecceitism" in Dusko Prelevic and Anand Vaidya (eds.), *The Epistemology of Modality and Philosophical Methodology*, Routledge. Chapter **6** has been published as "Grounding, Conceivability, and the Mind-Body Problem" in *Synthese* 195 (2):919-926 (2018), doi:10.1007/s11229-016-1254-2. Chapter **11** has been published as "Modal Ω-Logic: Automata, Neo-Logicism, and Set-Theoretic Realism" in Don Berkich and Matteo Vincenzo d'Alfonso (eds.), *On the Cognitive, Ethical, and Scientific Dimensions of Artificial Intelligence – Themes from IACAP 2016*, Springer (2019).

Table of Contents

6

Part IV: Epistemic Modality and the Philosophy of Mathematics

7

8

9

Chapter 1

Methodological Forward

This book concerns the foundations of epistemic modality. The work aims to advance our present understanding of the defining contours of epistemic modal space. I endeavor, then, to develop the theory of epistemic modality, by accounting for its interaction with metaphysical modality; deontic modality; the types of mathematical modality; the epistemic status of undecidable propositions and abstraction principles in the philosophy of mathematics; the apriori-aposteriori distinction; the modal profile of rational propositional intuition; and the types of intention, when the latter is interpreted as a modal mental state. In each chapter, I examine the philosophical significance of the foregoing, by demonstrating its import to a number of previously intransigent philosophical issues.

In Section **1**, I provide a summary of each of the chapters. In Section **2**, I examine the limits of competing proposals in the literature, and outline the need for a new approach.

1.1 Chapter Summary

In Chapter **2**, I aim to vindicate the thesis that cognitive computational properties are abstract objects implemented in physical systems. I avail of the equivalence relations countenanced in Homotopy Type Theory, in order to specify an abstraction principle for intensional, computational properties. The homotopic abstraction principle for intensional mental functions provides an epistemic conduit into our knowledge of cognitive algorithms as abstract objects. I examine, then, how intensional functions in Epistemic Modal Al-

gebra are deployed as core models in the philosophy of mind, Bayesian perceptual psychology, the program of natural language semantics in linguistics, and in quantum information theory, and I argue that this provides abductive support for the truth of homotopic abstraction. Epistemic modality can thus be shown to be both a compelling and a materially adequate candidate for the fundamental structure of mental representational states, comprising a fragment of the Language of Thought.

In Chapter **3**, I provide a mathematically tractable background against which to model both modal cognitivism and modal expressivism. I argue that epistemic modal algebras comprise a materially adequate fragment of the language of thought. I demonstrate, then, how modal expressivism can be regimented by modal coalgebraic automata, to which the above epistemic modal algebras are dually isomorphic. I examine, in particular, the virtues unique to the modal expressivist approach here proffered in the setting of the foundations of mathematics, by contrast to competing approaches based upon both the inferentialist approach to concept-individuation and the codification of speech acts via intensional semantics.

In Chapter **4**, I endeavor to establish foundations for the interaction between hyperintensional semantics and two-dimensional indexing. I examine the significance of the semantics, by developing three, novel interpretations of the framework. The first interpretation provides a characterization of the distinction between fundamental and derivative truths. The interaction between the hyperintensional truthmaker semantics and modal ontology is further examined. The second interpretation demonstrates how the elements of decision theory are definable within the semantics, and provides a novel account of the interaction between probability measures and hyperintensional grounds. The third interpretation concerns the contents of the types of intentional action, and the semantics is shown to resolve a puzzle concerning the role of intention in action.

In Chapter **5**, I aim further to redress the contention that epistemic possibility cannot be a guide to the principles of modal metaphysics. I argue that the interaction between the two-dimensional intensional framework and intensional plural quantification enables epistemic possibilities to target the haecceitistic properties of individuals. I outline the elements of plural logic, and I specify, then, a two-dimensional intensional formula encoding the relation between the epistemic possibility of haecceity comprehension and its metaphysical possibility. I examine the Julius Caesar problem as a test case. I conclude by addressing objections from the indeterminacy of ontological

principles relative to the space of epistemic possibilities, and from the consistency of epistemic modal space.

In Chapter **6**, I argue that Chalmers's (1996; 2010) two-dimensional conceivability argument against the derivation of phenomenal truths from physical truths risks being obviated by a hyperintensional regimentation of the ontology of consciousness. The regimentation demonstrates how ontological dependencies between truths about consciousness and about physics cannot be witnessed by epistemic constraints, when the latter are recorded by the conceivability – i.e., the epistemic possibility – thereof. Generalizations and other aspects of the philosophical significance of the hyperintensional regimentation are further examined.

In Chapter **7**, I aim to provide a precise account of the manner in which the properties of phenomenal consciousness can be understood as possessing accuracy-conditions. The account is restricted to an examination of sensory phenomenal properties. I provide a critical discussion of an approach to phenomenal content, which avails of the epistemic interpretation of two-dimensional intensional semantics. I examine several general issues for the foregoing approach, which I argue adduce against the sufficiency of the account as a theory of phenomenal representation. By contrast, I go on to provide a formally tractable account of the manner in which phenomenal properties can precisely compose via a category-theoretic semantics. The version of category-theoretic semantics that I proffer is inspired by Russell's (op. cit.) and Pettigrew's (ms) seminal work on the approach. I demonstrate how the proposal is empirically adequate, by constitutively incorporating some of the results concerning the kinds of attention and models of Bayesian perceptual psychology; and I argue that this concatenation of phenomenal properties into propositional form, assessable for accuracy or veridicality, explains the nature of the correctness-conditions for phenomenal consciousness. The second, central aim of the chapter is to explain how the phenomenological effects of attention interact with phenomenal representations, and to examine how the formal properties of phenomenal content relate to the logic of knowledge. I argue (i) that the phenomenological effects of attention do not affect the satisfaction of phenomenal knowledge, and thus do not affect the factivity thereof, and (ii) that the value-altering effects of attention bear on phenomenal content only by targeting its component parts, in a manner similar to how changes in the semantic values of sub-sentential expressions might affect the truth-condition for the sentence that the component expressions comprise.

In Chapter **8**, I provide a novel account of iterated epistemic states. I ar-

gue that states of epistemic determinacy might be secured by countenancing self-knowledge on the model of fixed points in monadic second-order modal logic, i.e. the modal μ-calculus. Despite the epistemic indeterminacy witnessed by the invalidation of modal axiom 4 in the sorites paradox – i.e. the KK principle: $\Box\phi \to \Box\Box\phi$ – an epistemic interpretation of the Kripke functors of a μ-automaton permits the iterations of the transition functions to entrain a principled means by which to account for necessary conditions on self-knowledge.

In Chapter **9**, I endeavor to provide a precise formulation of the dependence relations between the apriori and aposteriori epistemic profiles of mental states, beyond enabling and evidential conditions. I pursue an intensional approach to countenancing the notion of apriority, and I argue that the finer-grained dependence relations between the apriori and aposteriori can be regimented using the logic of explanatory ground. I advance four methods for countenancing the interaction between the intensional apriori and the aposteriori. The first method is witnessed by Bayesian confirmation measures, where the probability operators figuring therein are defined on epistemically possible worlds. The second method concerns the individuation-conditions on intensions defined on epistemically possible worlds. The third method targets the notion of the modally factive apriori. The fourth method targets iterations of the grounding operator, variably interpreted as expressing epistemic and metaphysical explanatory dependence. I argue, then, that the framework can be applied, in order to resolve two puzzles concerning the contrasting etiology and epistemic profiles of mental states: (i) immunity to error through misidentification, and (ii) the precise interaction between epistemic modality and deontic modality, as a test case for the relations between the intensional apriori and explanatory method.

In Chapter **10**, I aim to provide modal foundations for mathematical platonism. I examine Hale and Wright's (2009) objections to the merits and need, in the defense of mathematical platonism and its epistemology, of the thesis of Necessitism. In response to Hale and Wright's objections to the role of epistemic and metaphysical modalities in providing justification for both the truth of abstraction principles and the success of mathematical predicate reference, I examine the Necessitist commitments of the abundant conception of properties endorsed by Hale and Wright and examined in Hale (2013); examine cardinality issues which arise depending on whether Necessitism is accepted at first- and higher-order; and demonstrate how a two-dimensional intensional approach to the epistemology of mathematics, augmented with

Necessitism, is consistent with Hale and Wright's notion of there being epistemic entitlement rationally to trust that abstraction principles are true. Epistemic and metaphysical modality may thus be shown to play a constitutive role in vindicating the reality of mathematical objects and truth, and in explaining our possible knowledge thereof.

In Chapter **11**, I examine the philosophical significance of Ω-logic in Zermelo-Fraenkel set theory with choice (ZFC). The dual isomorphism between algebra and coalgebra permits Boolean-valued algebraic models of ZFC to be interpreted as coalgebras. The modal profile of Ω-logical validity can then be countenanced within a coalgebraic logic, and Ω-logical validity can be defined via deterministic automata. I argue that the philosophical significance of the foregoing is two-fold. First, because the epistemic and modal profiles of Ω-logical validity correspond to those of second-order logical consequence, Ω-logical validity is genuinely logical, and thus vindicates a neologicist conception of mathematical truth in the set-theoretic multiverse. Second, the foregoing provides a modal-computational account of the interpretation of mathematical vocabulary, adducing in favor of a realist conception of the cumulative hierarchy of sets.

In Chapter **12**, I aim to contribute to the analysis of the nature of mathematical modality, and to the applications of the latter to unrestricted quantification and absolute decidability. Rather than countenancing the interpretational type of mathematical modality as a primitive, I argue that the interpretational type of mathematical modality is a species of epistemic modality. I argue, then, that the framework of two-dimensional intensional semantics ought to be applied to the mathematical setting. The framework permits of a formally precise account of the priority and relation between epistemic mathematical modality and metaphysical mathematical modality. The discrepancy between the modal systems governing the parameters in the two-dimensional intensional setting provides an explanation of the difference between the metaphysical possibility of absolute decidability and our knowledge thereof.

In Chapter **13**, I endeavor to define the concept of indefinite extensibility in the setting of category theory. I argue that the generative property of indefinite extensibility in the category-theoretic setting is identifiable with the Kripke functors of modal coalgebraic automata, where the automata model Grothendieck Universes and the functors are further inter-definable with the elementary embeddings of large cardinal axioms. The Kripke functors definable in Grothendieck universes are argued to account for the ontological

expansion effected by the elementary embeddings in the category of sets. By characterizing the modal profile of Ω-logical validity, and thus the generic invariance of mathematical truth, modal coalgebraic automata are further capable of capturing the notion of definiteness, in order to yield a non-circular definition of indefinite extensibility.

In Chapter **14**, I aim to provide a modal logic for rational intuition. Similarly to treatments of the property of knowledge in epistemic logic, I argue that rational intuition can be codified by a modal operator governed by the axioms of a dynamic provability logic, which embeds GL within the modal μ-calculus. If rational intuition is identical with cognitive phenomenal properties of representational states such as beliefs, then – via correspondence results between modal logic and first-order logic [cf. van Benthem (1983; 1984/2003)] – a precise translation can be provided between the notion of 'intuition-of', i.e., the cognitive phenomenal properties of thoughts whose contents can concern the axioms of mathematical languages, and the modal operators regimenting the notion of 'intuition-that'. I argue that intuition-that can further be shown to entrain conceptual elucidation, by way of figuring as a dynamic-interpretational modality which induces the reinterpretation of domains of quantification (cf. Fine, 2005; 2006). Modalized rational intuition is therefore expressively equivalent to – and can crucially serve as a guide to the interpretation of – the entities, such as mathematical concepts, that are formalizable in monadic first- and second-order formal languages.

In Chapter **15**, I target a series of potential issues for a modal resolution to the alethic paradoxes. I aim, then, to provide a novel, epistemicist treatment of the alethic paradoxes. In response to Curry's paradox, the epistemicist solution that I advance enables the retention of both classical logic and the traditional rules for the alethic predicate: truth-elimination and truth-introduction. By availing of epistemic modal logic, the epistemicist approach permits, further, of a descriptively adequate explanation of the indeterminacy that is exhibited by epistemic states concerning liar-paradoxical sentences.

In Chapter **16**, I argue that the types of intention can be modeled as a type of modal operator. I delineate the intensional-semantic profiles of the types of intention, and provide a precise account of how the types of intention are unified in virtue of both their operations in a single, encompassing, epistemic modal space, and their role in practical reasoning. I endeavor to provide reasons adducing against the proposal that the types of intention are reducible to the mental states of belief and desire, where the former state

is codified by subjective probability measures and the latter is codified by a utility function. I argue, instead, that each of the types of intention – i.e., intention-in-action, intention-as-explanation, and intention-for-the-future – has as its aim the value of an outcome of the agent's action, as derived by her partial beliefs and assignments of utility, and as codified by the value of expected utility in evidential decision theory.

1.2   The Need for a New Approach

The proposal that mental representations can be defined as possibilities relative to states of information dates at least back to Wittgenstein (1921/1974), although there are a number of precursors to the literature in the twentieth century.[1] While novel, the limits of these incipient proposals consists in that they are laconic with regard to the explanatory foundations of the general approach.

   Wittgenstein writes: 'A picture is a fact. / The fact that the elements of a picture are related to one another in a determinate way represents that things are related to one another in the same way. / Let us call this connexion of its elements the structure of the picture, and let us call the possibility of this structure the pictorial form of the picture. / Pictorial form is the possibility that things are related to one another in the same way as the elements of the picture. A logical picture of facts is a thought. / 'A state of affairs is thinkable': what this means is that we can picture it to ourselves. / The totality of

---

[1] For an examination of epistemic logic in, e.g., the late medieval period, see Boh (1993).

   For the role of logical, rather than epistemic, modality in defining the modes of judgment, see Buridan (2001: 5.6), Kant (1787/1998: A74/B99-A76/B101), and Bolzano (1810/2004: 15-16). For the synthetic apriori determination of which of the possible predicates comprising a disjoint union ought to be applied to objects – i.e., transcendental logic – see Kant (op. cit: A53/B77-A57/B81; A571/B599-A574/B602). Anticipating Kripke (1980: 56), Husserl (1929/1999: §6) refers, in a section heading and the discussion therein, to transcendental logic as pertaining to conditions on the 'contingent apriori'.

   For the role of possibilities in accounting for the nature of subjective probability measures, i.e., partial belief, see Bernoulli (1713/2006: 211), Wittgenstein (op. cit.: 4.464, 5.15-5.152), and Carnap (1945). Bernoulli (op. cit.) writes: 'Something is possible if it has even a very small part of certainty, impossible if it has none or infinitely little. Thus something that has 1/20 or 1/30 of certainty is possible'. For subjective interpretations of probability, see Pascal (1654/1959), Laplace (1774/1986), Boole (1854), Ramsey (1926/1960), de Finetti (1937/1964), and Koopman (1940). For the history of the development of the theory of subjective probability, see Daston (1988; 1994) and Joyce (2011).

true thoughts is a picture of the world. / A thought contains the possibility of the situation of which it is the thought. What is thinkable is possible too' (op. cit.: 2.141-2.151, 3-3.02). Wittgenstein notes, further, that 'The theory of knowledge is the philosophy of psychology' (4.1121), and inquires: 'Does not my study of sign-language correspond to the study of thought processes which philosophers held to be so essential to the philosophy of logic? Only they got entangled for the most part in unessential psychological investigations, and there is an analogous danger for my method' (op. cit.).[2] Chapters **2** and **3** endeavor, as noted, to argue for a similar thesis, to the effect that epistemic modality comprises a materially adequate fragment of the Language of Thought, i.e., the computational structure and semantic values of the mental representations countenanced in philosophy and cognitive science.

Modal analyses of the notions of apriority and of states of information broadly construed are further proffered in Russell (1919), Lewis (1923), and Peirce (1933). Russell (op. cit.: 345-346) contrasts the possible truth-value of a propositional function (i.e., open formula) given an assignment of values to the variables therein with an epistemic – what he refers to as the 'ordinary' – interpretation of the modal according to which 'when you say of a proposition that it is possible, you mean something like this: first of all it is implied that you do not know whether it is true or false, and I think it is implied; secondly, that it is one of a class of propositions, some of which are known to be true. When I say, e.g., 'It is possible that it may rain to-morrow' ... We mean partly that we do not know whether it will rain or whether it will not, but also that we do know that that is the sort of proposition that is quite apt to be true, that it is a value of a propositional function of which we know some value to be true' (op. cit.: 346). Lewis (op. cit.: 172) defines the apriority of the laws of mathematical languages as consisting in their being 'true in all possible worlds'. Peirce (op. cit.: §65) writes:

'[L]et me say that I use the word information to mean a state of knowledge, which may range from total ignorance of everything except the meanings of words up to omniscience; and by informational I mean relative to such a state of knowledge. Thus, by 'informationally possible,' I mean possible so far as we, or the persons considered, know. Then, the informationally possible is that which in a given information is not perfectly known not to be true. The informationally necessary is that which is perfectly known to be

---

[2]The remarks are anticipated in Wittgenstein [1979: 21/10/14, 5/11/14, 10/11/14, 12/11/14 (pp. 16, 24-29)].

true. The informationally contingent, which in the given information remains uncertain, that is, at once possible and unnecessary'.

The notion of epistemic modality was, finally, stipulated independently by Moore (c.1941-1942/1962) in his commonplace book. According to Moore, 'epistemic' possibilities include that 'Its possible that [for some individual, a: a] is [glad] right now [iff] [a] may be [glad]', where 'I know that he's not' contradicts' the foregoing sentence (op. cit.: 187). Another instance of an epistemic possibility is advanced – 'It's possible that I'm not sitting down right now' – and analyzed as: 'It's not certain that I am' or 'I don't know that I am" (184).

In the contemporary literature, there is a paucity of works devoted to the nature of epistemic modality and its relation to other modalities. Recent books and edited volumes which examine aspects of epistemic modality include Gendler and Hawthorne (2002); Yablo (2008); Gendler (2010; Egan and Weatherson (2011); and Chalmers (2012). While the present work is focused on the foundations and philosophical significance of the epistemic interpretation of modal logic and semantics, there is some convergence, as well, with the literature on the epistemology of modality, i.e., on the conditions under which knowledge of modal properties obtains. Chapters **4-5** accounts for the relation between both epistemic and metaphysical modalities. For the sake of completeness, a critical summary of the relevant literature is thus included below.

The Gendler and Hawthorne volume includes seminal contributions to the theory of the relationship between epistemic and metaphysical modality. By contrast, this book provides foundations for the nature of epistemic modality, when the modality concerns apriority and conceivability, as well as the logic of knowledge and belief; makes contributions to our understanding of the ontology of consciousness, by regimenting the ontology of consciousness using hyperintensional grounding operators; examines the nature and philosophical extensions of epistemic logic; and examines the relations between epistemic modality and the variety of other modalities (e.g., metaphysical, deontic, and mathematical modalities; the types of intention in the setting of evidential decision theory; and the apriori-aposteriori distinction).

The papers on modal epistemology in Yablo (2008) predominantly concern the relation between epistemic and metaphysical modalities, and, in particular, non-trivial conditions on modal error. Issues for the epistemic interpretation of two-dimensional intensional semantics are examined; e.g., the absence of conditions on ascertaining when an epistemic possibility is

actual. The discussion is similar, in scope, to the discussions in the Gendler and Hawthorne volume. This book aims to redress the limits mentioned in the foregoing, and to proffer the positive proposals delineated above.

The Egan and Weatherson volume is comprised of papers which predominantly analyze epistemic modals in the setting of natural language semantics. Only three papers in the volume target epistemic possibilities as imaginable or conceptual possibilities; those by Chalmers ('The Nature of Epistemic Space'), Jackson ('Possibilities for Representation and Credence'), and Yalcin ('Non-factualism about Epistemic Modality'). Chalmers' paper examines some principles governing epistemic space and its interaction with metaphysical modality, as well as Kaplan's paradox. As mentioned, the book endeavors, by contrast, further to explain the individuation-conditions on the terms and intensions defined in epistemic modal space; examines the interaction between epistemic modality and various other types of modality; and examines the role that epistemic modal logic plays in resolving the knowability paradox and the alethic paradoxes, as well as undecidable sentences in the philosophy of mathematics. Jackson's paper argues that conceptual possibilities and metaphysical possibilities ought to be defined within a single space, in order both to avoid cases in which a sentence is conceptually possible although metaphysically impossible and to secure the representational adequacy of conceptually possible terms. Yalcin's paper argues that epistemic modal sentences in natural language semantics mirror the structure of the beliefs of speakers. Epistemic mental states are taken, then, to be expressive rather than representational, because the communication of epistemic modal and interrogative updates on an informational background shared by speakers is not truth-conditional.[3] The present approach contrasts to the foregoing, by not taking the values of expressions in natural language semantics to be a guide to the nature of mental states (cf. Evans, 1982).

Chalmers (2012) provides a book-length examination of the scrutability of truth, and the apriori entailment relations between different types of truths. The rigidity of intensions is availed of, in order to explain the relation between epistemic modality and metaphysical modality. By contrast, the book aims to examine the interaction between mental imagery and propositional imagination (i.e. the individuation of the intensions defined in epistemic

---

[3]The view that subject matters, broadly construed, have the form of an interrogative update on a set of worlds is anticipated by Lewis (1988/1998) and further defended by Yablo (2014 and Yalcin (2016). See Chapter **3**, for further discussion.

modal space), as well as the interaction between epistemic possibilities and the apriori-aposteriori distinction. The book aims (i) to provide explanatory foundations for the bare, two-dimensional semantic regimentation, as well as (ii) to examine novel philosophical extensions of the framework.

Gendler (2010) is a rare, empirically informed study of the limits of representational capacities, when they target counterfactual assignments of values to variables in thought experiments – e.g., the conditions under which there might be resistance attending the states of imagining that fictional characters have variant value-theoretic properties – and when implicit biases and unconscious sub-doxastic states affect the veridicality conditions of one's beliefs. One crucial distinction between Gendler's approach and the one pursued in this chapter, however, is that the former does not examine the interaction between apriority, the imagination, and modal semantics and logic.

In the literature on modal epistemology, Hale (2013) argues that modal knowledge ought to be pursued via the epistemology of essential definitions which specify conditions on sortal membership. Apriori knowledge of essence is explained in virtue of knowledge of the purely general terms – embedding no singular terms – which figure in the definitions. Thus – by being purely general – the essential properties and the objects falling in their extension have necessary being. Aposteriori knowledge of essential definitions can be pursued via theoretical identity statements, yet, because the terms figuring therein are not purely general, both the essential properties and the objects falling in their extension have contingent being. As mentioned, the book redefines the extant proposals in the ontology of consciousness using hyperintensional grounding operators. The ground-theoretic interpretation of the ontology of consciousness, and an examination of the bearing of the latter for the relation between conceivability and metaphysical possibility is, as noted, examined in Chapter **5**. Hale's higher-order Necessitist proposal is examined in further detail, in Chapter **9**.

Nichols (2006) features three essays on modal epistemology. Nichols' 'Imaginative Blocks and Impossibility' examines introspection-based tasks in developmental psychology, in order to account for the interaction between imaginative exercises and counterfactual judgments. Hill's 'Modality, Modal Epistemology, and the Metaphysics of Consciousness' examines the interaction between conceptual and metaphysical possibility, where conceptual possibilities are construed as Fregean thoughts, and the relation between conceivability and metaphysical possibility is then analyzed as the relation between Fregean thoughts (augmented by satisfaction-conditions such as

conceptual coherence) and empirical propositions. Sorensen's paper, 'Meta-conceivability and Thought Experiments', argues that meta-conceivable thought experiments are distinct from both conscious perceptual states and conceivable possibilities. [Sorensen (1999) argues that (thought) experiments track the consequences of reassignments of values to variables.] My approach differs from Hill's by arguing in favor of a possible worlds semantics for thoughts, which is able to recover the virtues attending the Fregean model, as well as in accounting for the relations between epistemic modal algebras and various other interpretations of modality, including the mathematical (cf. Chapter **9** and **11** for further discussion). My approach is similar in methodology to Nichols', although it endeavors to provide a naturalistically adequate explanation of the relationship between mental imagery (as a species of phenomenal property type) and the propositions entertained in imaginative exercises, i.e. the possibilities comprising epistemic modal space. Finally, my approach is similar to Sorensen's, in targeting both a formal semantic analysis of epistemic and related modalities, as well as the operators of knowledge and belief in the setting of epistemic modal logic.

Finally, a class of views in the epistemology of modality can be characterized as being broadly empiricist. Stalnaker (2003) and Williamson (2007; 2013) refrain from countenancing the notion of epistemically possible worlds; and argue instead either that the imagination is identified with cognitive processes taking the form of counterfactual presupposition (Williamson, 2007); that one's choice of the axioms governing modal logic should satisfy abductive criteria on theory choice (Williamson, 2013); or that metaphysical modalities are properties of the actual world (Stalnaker, op. cit.). Vetter (2013) argues for a reduction of modal notions to actual dispositional properties, and Roca-Royes (2016) pursues a corresponding modal empiricist approach, according to which knowledge of the *de re* possibilities of objects consists in the extrapolation of properties from acquaintance with objects in one's surround to formally similar objects, related by reflexivity and symmetry. Generally, according to the foregoing approaches, the method of modal epistemology proceeds by discerning the modal truths – captured, e.g., by abductively preferred theorems in modal logic; conditional propositions; and dispositional and counterfactual properties – and then working backward to the exigent incompleteness of an individual's epistemic states concerning such truths. By contrast, the approach advanced in this work both retains and provides explanatory foundations for epistemic modal space, and augments the examination by empirical research and an abductive methodology.

The foregoing texts either examine epistemic modality via natural language semantics; restrict their examination to the interaction between conceivable possibilities and metaphysical possibilities; provide a two-dimensional regimentation of formulas, on an epistemic interpretation, without providing an explanation of the source of, and individuation-conditions on, epistemic possibilities; eschew of epistemic possibilities; provide a naturalistic approach to the analysis of the imagination, without drawing on formal methods; or provide a formal analysis of the imagination, without drawing on empirical results.

The book endeavors, by contrast, to examine the interaction between epistemic modality and the computational theory of mind; metaphysical modality; deontic modality; the types of mathematical modality; the apriori-aposteriori distinction; the modal profile of rational intuition; and the types of intention, when the latter is interpreted as a modal mental state.

The model structures developed here are of interest in their own right. However, this work is principally concerned with, and examines, their philosophical significance, as witnessed by the new distinctions and relations that they induce. Beyond conditions on theoretical creativity, both formal regimentation and empirical confirmation are the best methods available for truth-apt philosophical inquiry into the constitutive role of computational and phenomenal properties in comprising the space of epistemic modality and the imagination; and the multiple points of convergence between apriori and epistemically possible truth and the most general, fundamental structure of the actual metaphysical world.

Part I: A Framework for Epistemic Modality

Chapter 2

Cognitivism about Epistemic Modality

This essay aims to vindicate the thesis that cognitive computational properties are abstract objects implemented in physical systems.[1] A recent approach to the foundations of mathematics is Homotopy Type Theory.[2] In Homotopy Type Theory, homotopies can be defined as equivalence relations on intensional functions. In this essay, I argue that homotopies can thereby figure in abstraction principles for intensional, cognitive computational properties.[3] Homotopies for intensional functions thus comprise identity criteria for some cognitive mechanisms. The philosophical significance of the foregoing is twofold. First, the proposal demonstrates how epistemic modality is a viable candidate for a fragment of the Language of Thought.[4] The identity of intensional functions and epistemic modal operators can be witnessed via a unique, epistemic interpretation of the algebraic semantics of modal logic. Second, the proposal serves to delineate one conduit for our epistemic access

---

[1]Cf. Turing (1950); Putnam (1967); Newell (1973); Fodor (1975); and Pylyshyn (1978).

[2]Cf. The Univalent Foundations Program (2013).

[3]For the first proposal to the effect that abstraction principles can be used to define abstracta such as cardinal number, see Frege (1884/1980: 68; 1893/2013: 20). For the locus classicus of the contemporary abstractionist program, see Hale and Wright (2001).

[4]See, by contrast, Stalnaker (1984), who argues that possible worlds model theory ought to model the pragmatics of intentional states, rather than internal representations of agents, and who argues against the existence of epistemic possibilities (2003: ch. 11). Given a metalanguage, a precedent to the current approach – which models thoughts and internal representations via possible worlds model theory – can be found in Wittgenstein (1921/1974: 2.15-2.151, 3-3.02).

to cognitive algorithms as abstract objects.[5]

In Section **2**, I provide an abstraction principle for cognitive algorithms, by availing of the equivalence relations countenanced in Homotopy Type Theory. In Section **3**, I define a topological boolean algebraic semantics for modal logic, where the modalities defined on the algebra range over states of information.[6] In Section **4**, I describe how models of Epistemic Modal Algebra are

---

[5]Bealer (1982) proffers a *non-modal* algebraic logic for intensional entities – i.e., properties, relations, and propositions – which avails of a λ-definable variable-binding abstraction operator (op. cit.: 46-48, 209-210). Bealer reduces modal notions to logically necessary conditions-cum-properties, as defined in his non-modal algebraic logic (207-209). The present approach differs from the foregoing by: (i) countenancing a *modal* algebra, on an epistemic interpretation thereof; (ii) treating the abstraction operator as a Fregean function from concepts to objects, rather than as a λ-operator; (iii) availing of the univalence axiom in Homotopy Type Theory – which collapses identity and isomorphism – in order to provide an equivalence relation for the abstraction principle pertinent to (ii); and (iv) demonstrating how the model is availed of in various branches of the cognitive sciences, such that Epistemic Modal Algebra may be considered a viable candidate for the Language of Thought.

Katz (1998) proffers a view of the epistemology of abstracta, according to which the propositions of the Language of Thought are abstract, and on which the syntax and the semantics for the propositions are innate (35). Katz suggests that the proposal is consistent with both a Fregean approach to propositions, according to which they are thoughts formed by the concatenation of concepts, and a Russellian approach, according to which they are structured tuples of non-conceptual entities (36). He endorses, ultimately, an approach to propositions which eschews Frege's (1892/1997) distinction between sense and reference, yet according to which they are 'abstract senses', or thoughts, correlated to natural language sentence types (114-115). That propositions are abstract is argued, then, to suffice for knowledge of abstract entities (op. cit.). One difference between Katz's proposal and the one here presented is that Katz rejects modal approaches to propositions, because the latter cannot countenance reductio proofs based on counterfactuals with impossible antecedents (38; cf. Lewis, 1973: I.6). Following, Lewis (op. cit.), the present approach does not avail of impossible worlds, i.e., worlds at which true contradictions obtain; and thus counterfactuals with impossible antecedents are vacuously true. If so, then Katz's argument against modal approaches to propositions can be circumvented. A second difference is that, on Katz's approach, the necessity of mathematical truths is argued to consist in reductio proofs, such that the relevant formulas will be true on all interpretations, and thus true of logical necessity (39). However, the endeavor to reduce the necessity of mathematical truths to the necessity of logical consequence would result in the preclusion, both of cases of informal proofs in mathematics, which can, e.g., involve diagrams (cf. Azzouni, 2004; Giaquinto, 2008: 1.2), and of mathematical truths which obtain in axiomatizable, yet non-logical mathematical languages such as Euclidean geometry. Finally, Katz rejects abstraction principles, and thus implicit definitions for abstract objects (105-106).

[6]The epistemic modality can be interpreted so as to target notions of conceivability

---

26

availed of when perceptual representational states are modeled in Bayesian perceptual psychology; when speech acts are modeled in natural language semantics; when transformations are defined in quantum information theory; and when knowledge, belief, intentional action, and rational intuition are modeled in philosophical approaches to the nature of propositional attitudes. This provides abductive support for the claim that Epistemic Modal Algebra is both a compelling and materially adequate candidate for a fragment of the Language of Thought. In Section **5**, I argue that the proposal (i) resolves objections to the relevant abstraction principles advanced by both Dean (2016) and Linnebo and Pettigrew (2014), and (ii) reveals a commitment to *in re* platonism in the appeal to cognitive mechanisms in Azzouni's (2004; forthcoming) version of nominalism. Section **6** provides concluding remarks.

2.1 An Abstraction Principle for Cognitive Algorithms

In this section, I specify a homotopic abstraction principle for intensional functions. Intensional isomorphism, as a jointly necessary and sufficient condition for the identity of intensions, is first proposed in Carnap (1947: §14). The isomorphism of two intensional structures is argued to consist in their logical, or L-, equivalence, where logical equivalence is co-extensive with the notions of both analyticity (§2) and synonymy (§15). Carnap writes that: '[A]n expression in S is L-equivalent to an expression in S' if and only if the semantic rules of S and S' together, without the use of any knowledge about (extra-linguistic) facts, suffice to show that the two have the same extension' (p. 56), where semantic rules specify the intended interpretation of the constants and predicates of the languages (4).[7] The current approach differs from Carnap's by defining the equivalence relation necessary for an abstraction principle for epistemic intensions on Voevodsky's (2006) Univalence Axiom, which collapses identity with isomorphism in the setting of

---

($\diamond$) and apriority ($\square \equiv \neg\diamond\neg$); belief and knowledge; and psychophysical possibilities. See Section **4**, for further discussion.

[7]For criticism of Carnap's account of intensional isomorphism, based on Carnap's (1937: 17) 'Principle of Tolerance' to the effect that pragmatic desiderata are a permissible constraint on one's choice of logic, see Church (1954: 66-67).

intensional type theory.[8] In the following section, I define, then, a class of models for Epistemic Modal Algebra.

Topological Semantics

In the topological semantics for modal logic, a frame is comprised of a set of points in topological space, a domain of propositions, and an accessibility relation:

$F = \langle X,R \rangle$;

$X = (X_x)_{x \in X}$; and

$R = (Rxy)_{x,y \in X}$ iff $R_x \subseteq X_x$ x X_x, s.t. if Rxy, then $\exists o \subseteq X$, with $x \in o$ s.t. $\forall y \in o(Rxy)$,

where the set of points accessible from a privileged node in the space is said to be open.[9] A model defined over the frame is a tuple, $M = \langle F,V \rangle$, with V a valuation function from subsets of points in F to propositonal variables taking the values 0 or 1. Necessity is interpreted as an interiority operator on the space:

$M,x \Vdash \Box\phi$ iff $\exists o \subseteq X$, with $x \in o$, such that $\forall y \in o$ $M,y \Vdash \phi$.

Homotopy Theory

Homotopy Theory countenances the following identity, inversion, and concatenation morphisms, which are identified as continuous paths in the topology. The formal clauses, in the remainder of this section, evince how homotopic morphisms satisfy the properties of an equivalence relation.[10]

---

[8]Note further that, by contrast to Carnap's approach, epistemic intensions are here distinguished from linguistic intensions (cf. Chapter 4, for further discussion), and the current work examines the philosophical significance of the convergence between epistemic intensions and formal, rather than natural, languages. For a translation from type theory to set theory – which is of interest to, inter alia, the definability of epistemic intensions in the setting of set theory (cf. Chapters **10** through **12**, below) – see Linnebo and Rayo (2012). For topological Boolean-valued models of epistemic set theory – i.e., a variant of ZF with the axioms augmented by epistemic modal operators interpreted as informal provability and having a background logic satisfying S4 – see Scedrov (1985), Flagg (1985), and Goodman (1990).

[9]In order to ensure that the Kripke semantics matches the topological semantics, X must further be Alexandrov; i.e., closed under arbitrary unions and intersections.

[10]The definitions and proofs at issue can be found in the Univalent Foundations Program (op. cit.: ch. 2).

p : [0,1] → X, with p(0) = x and p(1) = y;
f : $X_1 \to X_2$;
g : $X_1 \to X_2$;
H : X_1 x {0,1} → X_2, $H_{x,0}$ = f(x) and $H_{x,1}$ = g(x).

Reflexivity

$\forall$x,y:A$\forall$p(p : x $=_A$ y) : τ(x,y,p), with A and τ designating types, and U
designating a universe of elements, e:
$\forall\alpha$:A$\exists$e(α) : τ(α, α, $\mathtt{ref}_\alpha$);
p,q : (x $=_A$ y)
$\exists$r$\in$e : p $=_{(x=_A y)}$ q
$\exists\mu$: r $=_{(p=_{(x=_A y)}q)}$ s.

The Induction Principle

If:
$\forall$x,y:A$\forall$p(p : x $=_A$ y)$\exists\tau[\tau$(x,y,p)] $\wedge$ $\forall\alpha$:A$\exists$e(α) : τ(α, α, $\mathtt{ref}_\alpha$)
Then:
$\forall$x,y:A$\exists$p(p : x $=_A$ y)$\exists$e[$\mathtt{ind}_{=_A}$(τ,e,x,y,p) : τ(x,y,p), such that
$\mathtt{ind}_{=_A}$(τ,e,α,α,$\mathtt{ref}_\alpha$) $\equiv$ e(α)].

Symmetry

$\forall$A$\forall$x,y:A$\exists$H$_\Sigma$(x=y → y=x)
H$_\Sigma$:= p $\mapsto$ p^{-1}, such that
$\forall$x:A($\mathtt{ref}_x$ $\equiv$ $\mathtt{ref}_x^{-1}$).

Transitivity

$\forall$A$\forall$x,y:A$\exists$H$_T$(x=y → y=z → x=z)
H$_T$:= p $\mapsto$ q $\mapsto$ p $\bullet$ q, such that
$\forall$x:A[$\mathtt{ref}_x$ $\bullet$ $\mathtt{ref}_x$ $\equiv$ $\mathtt{ref}_x$].

Homotopic Abstraction

For all type families A,B, there is a homotopy:

H := $[(f \sim g) :\equiv \prod_{x:A}(f(x) = g(x))]$, where
$\prod_{f:A \to B}[(f \sim f) \wedge (f \sim g \to g \sim f) \wedge (f \sim g \to g \sim h \to f \sim h)]$,
such that, via Voevodsky's (2006) Univalence Axiom, for all type families
A,B:U, there is a function:
`idtoeqv` : $(A =_U B) \to (A \simeq B)$,
which is itself an equivalence relation:
$(A =_U B) \simeq (A \simeq B)$.

Abstraction principles for intensional computational properties take, then, the form:

- $\exists f,g[f(x) = g(x)] \simeq [f(x) \simeq g(x)]$.

2.2   Epistemic Modal Algebra

In Epistemic Modal Algebra, the topological boolean algebra, *A*, can be formed by taking the powerset of the topological space, X, defined above; i.e., $A = P(X)$. The domain of *A* is comprised of formula-terms – eliding propositions with names – assigned to elements of P(X), where the proposition-letters are interpreted as encoding states of information. The top element of the algebra is denoted '1' and the bottom element is denoted '0'. We interpret modal operators, f(x), – i.e., intensional functions in the algebra – as both concerning topological interiority, as well as reflecting *epistemic possibilities*. An Epistemic Modal-valued Algebraic structure has the form, F = $\langle A, D_{P(X)}, \rho \rangle$, where ρ is a mapping from points in the topological space to elements or regions of the algebraic structure; i.e., $\rho : D_{P(X)}$ x $D_{P(X)} \to A$. A model over the Epistemic-Modal Topological Boolean Algebraic structure has the form M = $\langle F, V \rangle$, where $V(a) \leq \rho(a)$ and $V(a,b) \wedge \rho(a, b) \leq V(b)$.[11] For all $x_{x/a,\phi}$,y$\in A$:
$\quad$ f(1) = 1;
$\quad$ f(x) $\leq$ x;
$\quad$ f(x $\wedge$ y) = f(x) $\wedge$ f(y);
$\quad$ f[f(x)] = f(x);
$\quad$ V(a, a) > 0;
$\quad$ V(a, a) = 1;

---

[11]See Lando (2015); McKinsey and Tarski (1944); and Rasiowa (1963), for further details.

V(a, b) = V(b, a);
V(a, b) ∧ V(b, c) ≤ V(a, c);
V(a = a) = ρ(a, a);
V(a, b) ≤ f[V(a, b)];
V($\neg\phi$) = ρ($\neg\phi$) – f(ϕ);
V($\diamond\phi$) = $\rho\phi$ – f[– V(ϕ)];
V($\Box\phi$) = f[V(ϕ)].

2.3 Examples in Philosophy, Cognitive Science, and Quantum Information Theory

The material adequacy of epistemic modal algebras as a fragment of the the Language of Thought is witnessed by the prevalence of possible worlds semantics – the model theory for which is algebraic (cf. Blackburn et al., 2001: ch. 5) – in cognitive psychology, and the convergences thereof with theoretical physics. Possible worlds model theory is availed of in the computational theory of mind, Bayesian perceptual psychology, natural language semantics, and the theory of quantum information.

Marcus (2001) argues that mental representations can be treated as algebraic rules characterizing the computation of operations on variables, where the values of a target domain for the variables are universally quantified over and the function is one-one, mapping a number of inputs to an equivalent number of outputs (35-36). Models of the above algebraic rules can be defined in both classical and weighted, connectionist systems: Both a single and multiple nodes can serve to represent the variables for a target domain (42-45). Temporal synchrony or dynamic variable-bindings are stored in short-term working memory (56-57), while information relevant to long-term variable-bindings are stored in registers (54-56). Examples of the foregoing algebraic rules on variable-binding include both the syntactic concatenation of morphemes and noun phrase reduplication in linguistics (37-39, 70-72), as well as learning algorithms (45-48). Conditions on variable-binding are further examined, including treating the binding relation between variables and values as tensor products – i.e., an application of a multiplicative axiom for variables and their values treated as vectors (53-54, 105-106). In order to account for recursively formed, complex representations, which he refers to as structured propositions, Marcus argues instead that the syntax

and semantics of such representations can be modeled via an ordered set of registers, which he refers to as 'treelets' (108).

A strengthened version of the algebraic rules on variable-binding can be accommodated in models of epistemic modal algebras, when the latter are augmented by cylindrifications, i.e., operators on the algebra simulating the treatment of quantification, and diagonal elements.[12] By contrast to Boolean Algebras with Operators, which are propositional, cylindric algebras define first-order logics. Intuitively, valuation assignments for first-order variables are, in cylindric modal logics, treated as possible worlds of the model, while existential and universal quantifiers are replaced by, respectively, possibility and necessity operators ($\diamond$ and $\square$) (Venema, 2013: 249). For first-order variables, $\{v_i \mid i < \alpha\}$ with α an arbitrary, fixed ordinal, $v_i = v_j$ is replaced by a modal constant $\mathbf{d}_{i,j}$ (op. cit: 250). The following clauses are valid, then, for a model, M, of cylindric modal logic, with $E_{i,j}$ a monadic predicate and T_i for i,j $< \alpha$ a dyadic predicate:

M,w $\Vdash$ p $\iff$ w$\in$V(p);

M,w $\Vdash$ $\mathbf{d}_{i,j}$ $\iff$ w$\in$E$_{i,j}$;

M,w $\Vdash$ $\diamond_i\psi$ $\iff$ there is a v with wT$_i$v and M,v $\vdash$ ψ (252).[13]

Finally, a cylindric modal algebra of dimension α is an algebra, $\mathbb{A} = \langle$A, +, $\bullet$, $-$, 0, 1, $\diamond_i$, $\mathbf{d}_{ij}\rangle_{i,j<\alpha}$, where $\diamond_i$ is a unary operator which is normal ($\diamond_i 0$ = 0) and additive [$\diamond_i$(x + y) = $\diamond_i$x + $\diamond_i$y)] (257).

The philosophical interest of cylindric modal algebras to Marcus' cognitive models of algebraic variable-binding is that variable substitution is treated in the modal algebras as a modal relation, while universal quantification is interpreted as necessitation. The interest of translating universal generalization into operations of epistemic necessitation is, finally, that – by

---

[12]See Henkin et al (op. cit.: 162-163) for the introduction of cylindric algebras, and for the axioms governing the cylindrification operators.

[13]Cylindric frames need further to satisfy the following axioms (op. cit.: 254):

1. p $\rightarrow$ $\diamond_i$p
2. p $\rightarrow$ $\square_i\diamond_i$p
3. $\diamond_i\diamond_i$p $\rightarrow$ $\diamond_i$p
4. $\diamond_i\diamond_j$p $\rightarrow$ $\diamond_j\diamond_i$p
5. $\mathbf{d}_{i,i}$
6. $\diamond_i(\mathbf{d}_{i,j} \wedge$ p) $\rightarrow$ $\square_i(\mathbf{d}_{i,j} \rightarrow$ p)

[Translating the diagonal element and cylindric (modal) operator into, respectively, monadic and dyadic predicates and universal quantification: $\forall$xyz[(T$_i$xy $\wedge$ E$_{i,j}$y $\wedge$ T$_i$xz $\wedge$ E$_{i,j}$z) $\rightarrow$ y = z] (op. cit.)]

7. $\mathbf{d}_{i,j}$ $\iff$ $\diamond_k(\mathbf{d}_{i,k} \wedge \mathbf{d}_{k,j})$.

identifying epistemic necessity with apriority – both the algebraic rules for variable-binding and the recursive formation of structured propositions can be seen as operations, the implicit knowledge of which is apriori.

In Bayesian perceptual psychology, the problem of underdetermination is resolved by availing of a gradational possible worlds model. The visual system is presented with a set of possibilities with regard, e.g., to the direction of a light source. So, for example, the direction of light might be originating from above, or it might be originating from below. The visual system computes the constancy, i.e. the likelihood that one of the possibilities is actual.[14] The computation of the perceptual constancy is an unconscious statistical inference, as anticipated by Helmholtz's (1878) conjecture.[15] The constancy places, then, a condition on the accuracy of the attribution of properties – such as boundedness and volume – to distal particulars.[16]

In the program of natural language semantics in empirical and philosophical linguistics, the common ground or 'context set' is the set of possibilities presupposed by a community of speakers.[17] Kratzer (1979: 121) refers to cases in which the above possibilities are epistemic as an 'epistemic conversational background', where the epistemic possibilities are a subset of objective or circumstantial possibilities (op. cit.). Modal operators are then defined on the space, encoding the effects of various speech acts in entraining updates on the context set.[18] So, e.g., assertion is argued to provide a truth-conditional update on the context set, whereas there are operator updates, the effects of which are not straightforwardly truth-conditional and whose semantic values must then be defined relative to an array of intensional parameters (including a context – agent, time, location, et al. – and a tuple of indices).

In quantum information theory, let a constructor be a computation defined over physical systems. Constructors entrain nomologically possible

---

[14]Cf. Mamassian et al. (2002).

[15]For the history of the integration of algorithms and computational modeling into contemporary visual psychology, see Johnson-Laird (2004).

[16]Cf. Burge (2010), and Rescorla (2013), for further discussion. A distinction ought to be drawn between unconscious perceptual representational states – as targeted in Burge (op. cit.) – and the inquiry into whether the properties of phenomenal consciousness have accuracy-conditions – where phenomenal properties are broadly construed, so as to include, e.g., color-phenomenal properties, as well as the property of being aware of one's perceptual states.

[17]Cf. Stalnaker (1978).

[18]Cf. Kratzer (op. cit.); Stalnaker (op. cit.); Lewis (1980); Heim (1992); Veltman (1996); von Fintel and Heim (2011); and Yalcin (2012).

transformations from admissible input states to output states (cf. Deutsch, 2013). On this approach, information is defined in terms of constructors, i.e., intensional computational properties. The foregoing transformations, as induced by constructors, are referred to as tasks. Because constructors encode the counterfactual to the effect that, were an initial state to be computed over, then the output state would result, modal notions are thus constitutive of the definition of the tasks at issue. There are, further, both topological and algebraic aspects of the foregoing modal approach to quantum computation.[19] The composition of tasks is formed by taking their union, where the union of tasks can be satisfiable while its component tasks might not be. Suppose, e.g., that the information states at issue concern the spin of a particle. A spin-state vector will be the sum of the probabilities that the particle is spinning either upward or downward. Suppose that there are two particles which can be spinning either upward or downward. Both particles can be spinning upward; spinning downward; particle-1 can be spinning upward while particle-2 spins downward; and vice versa. The state vector, V which records the foregoing possibilities – i.e., the superposition of the states – will be equal to the product of the spin-state of particle-1 and the spin-state of particle-2. If the particles are both spinning upward or both spinning downward, then V will be .5. However – relative to the value of each particle vector, referred to as its eigenvalue – the probability that particle-1 will be spinning upward is .5 and the probability that particle-2 will be spinning downward is .5, such that the probability that both will be spinning upward or downward = .5 x .5 = .25. Considered as the superposition of the two states, V will thus be unequal to the product of their eigenvalues, and is said to be entangled. If the indeterminacy evinced by entangled states is interpreted as inconsistency, then the computational properties at issue might further have to be hyperintensional.[20] Modal notions play, further, an ineliminable role in physical ontology and in the theoretical and observational characterizations thereof. In the interpretation of physical theories, Ruetsche (2011: 9) argues, further, that '[T]o interpret a physical theory is to characterize the worlds possible according to that theory'. She notes that the interpretation of a physical theory falls into two phases, both of

<hr />

[19]For an examination of the interaction between topos theory and an S4 modal axiomatization of computable functions, see Awodey et al. (2000).

[20]The nature of the indeterminacy in question is examined in Hawthorne (2010) and Khudairi, 'Entanglement, Modality, and Indeterminacy' (ms). For a thorough examination of approaches to the ontology of quantum mechanics, see Arntzenius (2012: ch. 3).

which constitutively involve modal notions. The first phase consists in the specification of the possible states of the models at issue; the possible values taken with regard to the observable kinematics thereof, e.g., the expectation values of Hilbert operators; and then the laws comprising the dynamic transformations therein (op. cit.). The second phase targets the nomologically possible worlds in which the properties of the first phase of interpretation are satisfied (op. cit.). In quantum field theory, unital algebras are availed of, as well, where the algebras are closed under commutativity and associativity; scalar multiplication by complex numbers, such that for all vectors A,B, and complex numbers c_n: $c_1(A + B) = c_1A + c_1B$; $(c_1 + c_2)A = c_1A + c_2A$; $c_1(c_2A) = (c_1c_2)A$; $(c_1A)B - A(c_1B) = c_1(AB)$; and multiplicative identity, i.e., for an element I, $AI = IA = A$ (74). Finally, Belot (2011) argues that – in the dispute between relationalist and substantivalist conceptions of spatial geometry – a relationalist proposal can be proffered, according to which – by analogy with nomological possibilities – the different metric relations that can be grounded in the intrinsic points comprising spatial regions are geometric possibilities.

Finally, Epistemic Modal Algebra, as a fragment of the Language of Thought, is able to delineate the fundamental structure of the propositional attitudes targeted in 20th century philosophy; notably knowledge, belief, intentional action, and rational propositional intuition. In Chapter **16**, I argue, e.g., that the types of intention – acting intentionally; referring to an intention as an explanation for one's course of action; and intending to pursue a course of action in the future – can be modeled as modal operators, whose semantic values are defined relative to an array of intensional parameters. Field (2001: 85-86) argues that – because possible worlds models of intentional states entrain logical equivalence for contents that ought to be distinct – such models underdetermine the identification of the intentions subserving agents' behavior . The objection can be answered, in virtue of there being unique arrays of intensional parameters relative to which each of the types of intention (modeled as modal operators) receive their value. E.g., an agent can be said to act intentionally iff her 'intention-in-action' receives a positive semantic value, where a necessary condition on the latter is that there is at least one world in her epistemic modal space at which – relative to a context of a particular time and location, which constrains the admissibility of her possible actions as defined at a first index, and which subsequently constrains the outcome thereof as defined at a second index – the intention is realized:

$[\![$Intenton-in-Action$(\phi)]\!]_w = 1$ only if $\exists$w'$[\![\phi]\!]^{w',c(=t,l),a,o} = 1$.

The agent's intention to pursue a course of action at a future time – i.e., her 'intention-for-the-future' – can receive a positive value only if there is a possible world and a future time, relative to which the possibility that a state, ϕ, is realized can be defined. Thus:

$[\![\text{Intention-for-the-future}(\phi)]\!]_w = 1$ only if $\exists w' \forall t \exists t'[t< t' \wedge [\![\phi]\!]^{w',t'} = 1]$.

In the setting of epistemic logic, epistemic necessity can further be modeled in a relational semantics encoding the property of knowledge, whereas epistemic possibility might encode the property of belief (cf. Hintikka, 1962; Fagin et al., 1995; Meyer and van der Hoek, 1995; Williamson, 2009). Finally, in Chapter **13**, I treat Gödel's (1953) conception of rational propositional intuition as a modal operator in the setting of a bimodal, dynamic provability logic, and demonstrates how – via correspondence theory – the notion of 'intuition-of', i.e. a property of awareness of one's cognitive states, can be shown to be formally equivalent to the notion of 'intuition-that', i.e. a modal operator concerning the value of the propositional state at issue.[21]

2.4   Objections and Replies

Dean (2016) raises two issues for a proposal similar to the foregoing, namely that algorithms – broadly construed – can be defined via abstraction principles which specify equivalence relations between implementations of computational properties in isomorphic machines.[22] Dean's candidate abstraction principle for algorithms as abstracts is: that the algorithm implemented by

---

[21]The correspondence results between modal propositional and first-order logic are advanced in van Benthem (1983; 1984/2003) and Janin and Walukiewicz (1996). Availing of correspondence theory in order to account for the relationship between the notions of 'intuition-of' and 'intuition-that' resolves an inquiry posed by Parsons (1993: 233). As a dynamic interpretational modality, rational intuition can further serve as a guide to possible reinterpretations both of quantifier domains (cf. Fine, 2005) and of the extensions of mathematical vocabulary such as the membership-relation (cf. Uzquiano, 2015). This provides an account of Gödel's (op. cit.; 1961) suggestion that rational intuition can serve as a guide to conceptual elucidation.

[22]Fodor (2000: 105, n.4) and Piccinini (2004) note that the identification of mental states with their functional roles ought to be distinguished from identifying those functional roles with abstract computations. Conversely, a computational theory of mind need not be committed to the identification of abstract, computational operations with the functional organization of a machine. Identifying abstract computational properties with the functional organization of a creature's mental states is thus a choice point, in theories of the nature of mental representation.

M_1 = the algorithm implemented by M_2 iff $M_1 \simeq M_2$.[23] Both issues target the uniqueness of the algorithm purported to be identified by the abstraction principle.

The first issue generalizes Benacerraf's (1965) contention that, in the reduction of number theory to set theory, there must be, and is not, a principled reason for which to prefer the identification of natural numbers with von Neumann ordinals (e.g., $2 = \{\emptyset,\{\emptyset\}\}$), rather than with Zermelo ordinals (i.e., order-types of well-orderings).[24] The issue is evinced by the choice of whether to define algorithms as isomorphic *iterations* of state transition functions (cf. Gurevich, 1999), or to define them as isomorphic *recursions* of functions which assign values to a partially ordered set of elements (cf. Moschovakis, op. cit.). Linnebo and Pettigrew (2014: 10) argue similarly that, for two 'non-rigid' structures which admit of non-trivial automorphisms, one can define a graph which belies their isomorphism. E.g., let an abstraction principle be defined for the isomorphism between S and S*, such that

$\forall$S,S*$[\mathbf{A}$S $= \mathbf{A}$S* iff $\langle$S, $R_1 \ldots R_n \rangle \simeq \langle$S*, $R^*_1 \ldots R^*_n \rangle]$.

However, if there is a graph, G, such that:

S $= \{v_1, v_2\}$, and R $= \{\langle v_1, v_2 \rangle, \langle v_2, v_1 \rangle\}$,

then one can define an automorphism, f : G $\simeq$ G, such that f$(v_1) = v_2$ and f$(v_2) = v_1$, such that S* $= \{v_1\}$ while R* $= \{\langle v^*_1, v^*_1 \rangle\}$. Then S* has one element via the automorphism, while S has two. So, S and S* are not, after all, isomorphic.

The second issue is that complexity is crucial to the identity criteria of algorithms. Two algorithms might be isomorphic, while the decidability of one algorithm is proportional to a deterministic *polynomial* function of the size of its input – with k a member of the natural numbers, N, and TIME referring to the relevant complexity class: $\bigcup_{k \in N}$TIME(n^k) – and the decidability of the second algorithm will be proportional to a deterministic *exponential* function of the size of its input – $\bigcup_{k \in N}$TIME(2^{n^k}). The deterministic polynomial time complexity class is a subclass of the deterministic exponential time complexity class. However, there are problems decidable by algorithms only in polynomial time (e.g., the problem of primality testing, such that, for any two natural numbers, the numbers possess a greatest common divisor equal to 1), and only in exponential time (familiarly from logic, e.g., the

---

[23]Cf. Moschovakis (1998).

[24]Cf. Zermelo (1908/1967) and von Neumann (1923/1967). Well-orderings are irreflexive, transitive, binary relations on all non-empty sets, which define a least or distinguished element in the sets.

problem of satisfiability – i.e., whether, for a given formula, there exists a model which can validate it – and the problem of validity – i.e. whether a satisfiable formula is valid).[25]

Both issues can be treated by noting that Dean's discussion targets abstraction principles for the very notion of a computable function, rather than for abstraction principles for cognitive computational properties. It is a virtue of homotopic abstraction principles for cognitive intensional functions that both the temporal complexity class to which the functions belong, and the applications of the model, are subject to variation. Variance in the cognitive roles, for which Epistemic Modal Algebra provides a model, will crucially bear on the nature of the representational properties unique to the interpretation of the intensional functions at issue. Thus, e.g., when the internal representations in the Language of Thought – as modeled by Epistemic Modal Algebra – subserve perceptual representational states, then their contents will be individuated by both the computational constancies at issue and the external, environmental properties – e.g., the properties of lightness and distance – of the perceiver.[26] A further virtue of the foregoing is that variance in the coding of Epistemic Modal Algebras – i.e. in the types of information over which the intensional functions will be defined – by contrast to a restriction of the Language of Thought to mathematical languages such as Peano arithmetic, permits homotopic abstraction principles to circumvent the Burali-Forti paradox for implicit definitions based on isomorphism.[27]

The examples of instances of Epistemic Modal Algebra – witnessed by the possible worlds models in Bayesian perceptual psychology, linguistics, quantum information theory, and philosophy of mind – provide abductive support for the existence of the intensional functions specified in homotopic abstraction principles. The philosophical significance of independent, abductive support for the existence of epistemic modalities in the philosophy of mind and cognitive science is that the latter permits a circumvention of the objections to the abstractionist foundations of number theory that have accrued since its contemporary founding (cf. Wright, 1983). Eklund (2006) suggests, e.g., that the existence of the abstract objects which are the refer-

---

[25]For further discussion, see Dean (2015).

[26]The computational properties at issue can also be defined over non-propositional information states, such as cognitive maps possessed of geometric rather than logical structure. See, e.g., O'Keefe and Nadel (1978); Camp (2007); and Rescorla (2009).

[27]Cf. Burali-Forti (1897/1967). Hodes (1984,a) and Hazen (1985) note that abstraction principles based on isomorphism with unrestricted comprehension entrain the paradox.

ents of numerical term-forming operators might need to be secured, prior to assuming that the abstraction principle for cardinal number is true. While Hale and Wright (2009) maintain, in response, that the truth of the relevant principles will be prior to the inquiry into whether the terms defined therein refer, they provide a preliminary endorsement of an 'abundant' conception of properties, according to which identifying the sense of a predicate will be sufficient for predicate reference.[28] One aspect of the significance of empirical and philosophical instances of models of Epistemic Modal Algebra is thus that, by providing independent, abductive support for the truth of the homotopic abstraction principles for cognitive algorithms, the proposal remains neutral on the status of 'sparse' versus 'abundant' conceptions of properties. Another aspect of the philosophical significance of possible worlds semantics being availed of in Bayesian vision science, empirical linguistics, and quantum information theory, is that it belies the purportedly naturalistic grounds for Quine's (1963/1976) scepticism of *de re* modality.

A final objection derives from the work of Azzouni (2004; forthcoming). Azzouni (op. cit.) argues for a position which he refers to as 'quantifier neutralism', according to which the truth-conditions for quantifiers are sufficient for identifying the latter, without requiring reference to domains. Thus, reference to the purported domains over which quantifiers range is a superfluous and misleading metaphor. On this approach, the truth-conditions for quantifiers will be theory-relative. In the absence of objects comprising domains over which the quantifiers are purported to range, there is thus no reason to believe that the truth-conditions in mathematical and scientific theories are ontologically-committing. He suggests, then, that the explanation of the belief that there are objects derives from the 'involuntary' perception thereof (forthcoming: 282). The experience as of objects is then attributed to 'internal cognitive mechanisms' (286; see also pp. 8-9, 264, and the remainder of section 7.5).[29]

Azzouni's suggestion that there are 'internal cognitive mechanisms' sub-

---

[28]For identity conditions on abundant properties – where the domain of properties, in the semantics of second-order logic, is a subset of the domain of objects, and the properties are definable in a metalanguage by predicates whose satisfaction-conditions have been fixed – see Hale (2013). For a generalization of the abundant conception, such that the domain of properties is isomorphic to the powerset of the domain of objects, see Cook (2014).

[29]For an examination of the psychological mechanisms necessary for object-perception, see, inter alia, Treisman and Gelade (1980); Spelke and Kinzler (2007); and Scholl and Flombaum (2010).

serving apparent object perception might betray an ontological commitment to there being abstracta; namely, the intensional, cognitive computational properties countenanced in Bayesian perceptual psychology. Azzouni (op. cit.: 7.8 - 7.9) anticipates and endeavors to counter the foregoing reply, by noting that the optimality of theories proffered in logic, metaphysics, and the natural sciences requires no commitment to the existence of the entities mentioned therein.

Azzouni's argument against the existence of objects – concreta and abstracta alike – hinges on the individuation-conditions for quantifiers. He argues, as noted, that objectual, first-order quantifiers can be defined via metalinguistic truth-conditions (op. cit.: 10). The metalinguistic truth-conditions take, e.g., the form: $\exists$x.(Px) iff $\exists$X.δX (op. cit.). However, the foregoing theory of quantification is unduly restrictive. For example, the truth-condition for probability quantifiers is given by the formula:

(Px $\geq$ r)[ϕx],

where x is a bound second-order variable; P(x) encodes the probability that the set {x:ϕx} is less than or equal to r; *r is an element of the intersection between a subset of hereditarily countable sets and the real interval, [0,1]*; and the quantifier, (Px $\geq$ r), is closed under finite conjunctions and disjunctions [cf. Keisler (1985: 509-510); Barwise and Feferman (1985: 507)]. Crucially, the metalinguistic truth-condition for probability quantifiers makes explicit reference to the real numbers.[30] The claim, then, that quantifiers can be defined without reference to domains of objects does not generalize. So, if ontological commitment is generally to be eschewed of, the argument must proceed via a different means.

2.5   Concluding Remarks

In this essay, the equivalence relations countenanced in Homotopy Type Theory were availed of, in order to specify an abstraction principle for intensional, computational properties. The homotopic abstraction principle for intensional mental functions provides an epistemic conduit into our knowledge of cognitive algorithms as abstract objects. Because intensional functions in

---

[30]Note that, by contrast to Field's (1980/2016: 55-56) approach, Azzouni's eschewal of ontological commitment to concrete objects would preclude redefining the reals as relations of betweeness and congruence on points of spacetime. For further discussion of individuation-conditions on concepts of the reals, see Chapter **4**.

Epistemic Modal Algebra are deployed as core models in the philosophy of mind, Bayesian visual psychology, natural language semantics, and quantum information theory, there is independent abductive support for the truth of homotopic abstraction. Epistemic modality may thereby be recognized as both a compelling and a materially adequate candidate for the fundamental structure of mental representational states, and as thus comprising a fragment of the Language of Thought.

Chapter 3

Modal Cognitivism and Modal Expressivism

This essay endeavors to reconcile two approaches to the intensional foundations of thought: modal cognitivism and modal expressivism. The novel contribution of the paper is its argument for a reconciliation between the two positions, by providing a hybrid account in which both internal cognitive architecture, on the model of epistemic possibilities, as well as modal automata, are accommodated, while retaining what is supposed to be their unique and inconsistent roles.

Modal cognitivism is the proposal that the internal representations comprising the language of thought can be modeled via a possible world semantics. Modal expressivism has, in turn, been delineated in two ways. On the first approach, the presuppositions shared by a community of speakers have been modeled as circumstantial possibilities (cf. Kratzer, 1979; Stalnaker, 1978, 1984). Speech acts have in turn been modeled as modal operators, the semantic values of which are then defined relative to an array of intensional parameters (Stalnaker, op. cit.; Veltman, 1996; Yalcin, 2007). On the second approach, the content of concepts is supposed to be individuated via the ability to draw inferences, and the pragmatic abilities of individuals have been modeled as automata comprised of two transition functions. A counterfactual transition functional – encoding the recognition of distal properties – determines the range of admissible values for another transition function encoding the individual's actions (cf. Brandom, 2008). Inferential conditions constitutive of concept possession are then taken to have the same counterfactual form as the foregoing functions (Brandom, 2014), while truth-evaluable de-

scriptions of the automata are specified in a metalanguage (Brandom, 2008). Both the modal approach to shared information and the speech acts which serve to update the latter, and the modal-inferential approach to concept-individuation – are thus consistent with mental states having semantic values or truth-conditional characterizations.[1]

So defined, the modal cognitivist and modal expressivist approaches have been assumed to be in constitutive opposition. While the cognitivist proposal avails of modal resources in order to model the internal representations comprising an abstract language of thought, the expressivist proposal targets informational properties which extend beyond the remit of internal cognitive architecture: both the form and the parameters relevant to determining the semantic values of linguistic utterances, where the informational common ground is taken to be reducible to circumstantial possibilities; and the individuation of the contents of concepts on the basis of inferential behavior.

In this paper, I provide a background mathematical theory, in order to account for the reconciliation of the cognitivist and expressivist proposals. I avail, in particular, of the dual isomorphism between Boolean-valued models of epistemic modal algebras and coalgebras; i.e., labeled transition systems defined in the setting of category theory.[2] The functors of coalgebras permit of flexible interpretations, such that they are able to characterize both modal logics as well as discrete-state automata. I argue that the correspondence be-

---

[1] The notions of cognitivism and expressivism here targeted concern the role of internal – rather than external – factors in countenancing the nature of thought and information (cf. Fodor, 1975; Haugeland, 1978). Possible worlds semantics is taken then to provide the most descriptively adequate means of countenancing the structure of the foregoing. Delineating cognitivism and expressivism by whether the positions avail of internal representations is thus orthogonal to the eponymous dispute between realists and antirealists with regard to whether mental states are truth-apt, i.e., have a representational function, rather than being non-representational and non-factive, even if real (cf. Dummett, 1959; Blackburn, 1984; Price, 2013). Whereas the type of modal cognitivism examined here assumes that thoughts and information take exclusively the form of internal representations, the target modal expressivist proposals assume that information states are exhaustively individuated by both linguistic behavior and conditions external to the cognitive architecture of agents.

[2] For an algebraic characterization of dynamic-epistemic logic, see Kurz and Palmigiano (2013). Baltag (2003) develops a colagebraic semantics for dynamic-epistemic logic, where coalgebraic functors are intended to record the informational dynamics of single- and multi-agent systems. The current approach differs from the foregoing by examining the dual isomorphism between static epistemic modal algebras and coalgebraic automata, where the functors in the latter are interpreted as expressive of the epistemic states comprising the modal algebras.

tween epistemic modal algebras and modal coalgebraic automata is sufficient then for the provision of a mathematically tractable, modal foundation for thought and action.[3]

In Section **2**, I provide the background mathematical theory, in order to account for the reconciliation of the cognitivist and expressivist proposals.

In Section **3**, I outline an expressivist semantics for epistemic modals. Modal coalgebraic automata are argued, finally, to be preferred as models of modal expressivism, by contrast to the speech-act and inferentialist approaches, in virtue of the advantages accruing to the model in the philosophy of mathematics. The interest in modal coalgebraic automata consists, in particular, in the range of mathematical properties that can be recovered on the basis thereof.[4] By contrast to the above competing approaches to modal expressivism, the functors of modal coalgebraic automata are able both to model and explain elementary embeddings in the category of sets; the intensions of mathematical terms; as well as the modal profile of Ω-logical consequence.

Section **4** provides concluding remarks.

---

[3]The proposal that possible worlds semantics comprises the model for thoughts and propositions is anticipated by Wittgenstein (1921/1974: 2.15-2.151, 3-3.02); Chalmers (2011); and Jackson (2011). Their approaches depart, however, from the one here examined in the following respects. (i) Wittgenstein (op. cit.: 1-1.1) has been interpreted as endorsing an identity theory of propositions, which does not distinguish between internal thoughts and external propositions (cf. McDowell, 1994: 27; and Hornsby, 1997: 1-3). How the identity theory of propositions is able to accommodate Wittgenstein's suggestion that a typed hierarchy of propositions can be generated – only if the class of propositions has a general form and the sense of propositions over which operations range is invariant by being individuated by the possibilities figuring as their truth and falsity conditions (cf. Wittgenstein, 1979: 21/11/16, 23/11/16, 7/11/17; and Potter, 2009: 283-285 for detailed discussion) – is an open question. Wittgenstein (1921/1974: 5.5561) writes that 'Hierarchies are and must be independent of reality', although provides no account of how the independence can be effected. (ii) Jackson (2008: 48-50) distinguishes between personal and subpersonal theories by the role of neural science in individuating representational states (cf. Shea, 2013, for further discussion), and argues in favor of a 'personal-level implicit theory' for the possible worlds semantics of mental representations. (iii) Chalmers (2012: 462-463) argues for a hybrid cognitivist-expressivist approach, according to which epistemic intensions – i.e. functions from epistemically possible worlds to extensions – are individuated by their inferential roles.

[4]See Wittgenstein (2001: IV, 4-6, 11, 30-31), for a prescient expressivist approach to the modal profile of mathematical formulas.

3.1   The Hybrid Proposal

3.1.1   Epistemic Modal Algebra

In Epistemic Modal Algebra, a topological Boolean algebra, A, can be formed by taking the powerset of a topological space, X; $A = $ P(X). The domain of A is comprised of formula-terms – eliding propositions with names – assigned to elements of P(X), where the formula-terms encode *epistemic possibilities*; i.e., possibilities defined relative to states of information. The top element of the algebra is denoted '1' and the bottom element is denoted '0'. We interpret the modal operators, f(x), – i.e., monotonic intensional functions in the algebra – as both equivalent to concepts, as well as countenancing the property of topological interiority. An Epistemic Modal-valued Algebraic structure has the form, F = $\langle A,$ $D_{P(X)},$ $\rho \rangle$, where ρ is a mapping from points in the topological space to elements or regions of the algebraic structure; i.e., $\rho :$ $D_{P(X)}$ x $D_{P(X)} \to A$. A model over the Epistemic-Modal Topological Boolean Algebraic structure has the form M = $\langle$F, V$\rangle$, where V(a) $\leq \rho$(a) and V(a,b) $\wedge \rho$(a, b) $\leq$ V(b).[5] For all $x_{x/a,\phi}$,y$\in A$:

    f(1) = 1;
    f(x) $\leq$ x;
    f(x $\wedge$ y) = f(x) $\wedge$ f(y);
    f[f(x)] = f(x);
    V(a, a) $>$ 0;
    V(a, a) = 1;
    V(a, b) = V(b, a);
    V(a, b) $\wedge$ V(b, c) $\leq$ V(a, c);
    V(a = a) = ρ(a, a);
    V(a, b) $\leq$ f[V(a, b)];
    V($\neg\phi$) = $\rho(\neg\phi) -$ f(ϕ);
    V($\diamond\phi$) = $\rho\phi -$ f[$-$ V(ϕ)];
    V($\Box\phi$) = f[V(ϕ)] (cf. Lando, op. cit.).
    An abstraction principle for epistemic intensions can be defined as follows:
    For all types, A,B, there is a homotopy:

H := [(f $\sim$ g) :$\equiv \prod_{x:A}$(f(x) = g(x)], where
$\prod_{f:A \to B}$[(f $\sim$ f) $\wedge$ (f $\sim$ g $\to$ g $\sim$ f) $\wedge$ (f $\sim$ g $\to$ g $\sim$ h $\to$ f $\sim$ h)],

   [5]See Lando (2015); McKinsey and Tarski (1944); and Rasiowa (1963), for further details.

such that, via Voevodsky's (2006) Univalence Axiom, for all type families A,B:U, there is a function:
`idtoeqv` : $(A =_U B) \to (A \simeq B)$,
which is itself an equivalence relation:
$(A =_U B) \simeq (A \simeq B)$.

Abstraction principles for intensional computational properties take, then, the form:

- $\exists f,g[f(x) = g(x)] \simeq [f(x) \simeq g(x)]$.

3.1.2 Modal Coalgebraic Automata

Modal coalgebraic automata can be thus characterized. Let a category C be comprised of a class Ob(C) of objects and a family of arrows for each pair of objects C(A,B) (Venema, 2007: 421). A functor from a category C to a category D, $\mathbf{E}$: C $\to$ D, is an operation mapping objects and arrows of C to objects and arrows of D (422). An endofunctor on C is a functor, $\mathbf{E}$: C $\to$ C (op. cit.).

A $\mathbf{E}$-coalgebra is a pair $\mathbb{A} = (A, \mu)$, with A an object of C referred to as the carrier of $\mathbb{A}$, and μ: A $\to \mathbf{E}(A)$ is an arrow in C, referred to as the transition map of $\mathbb{A}$ (390).

As, further, a coalgebraic model of modal logic, $\mathbb{A}$ can be defined as follows (407):

For a set of formulas, Φ, let $\nabla\Phi := \Box \bigvee \Phi \land \bigwedge \Diamond\Phi$, where $\Diamond\Phi$ denotes the set $\{\Diamond\phi \mid \phi \in \Phi\}$ (op. cit.). Then,

$\Diamond\phi \equiv \nabla\{\phi, \mathrm{T}\}$,

$\Box\phi \equiv \nabla\varnothing \lor \nabla\phi$ (op. cit.).

Let, then, an $\mathbf{E}$-coalgebraic modal model, $\mathbb{A} = \langle S,\lambda,R[.]\rangle$, be such that $\mathbb{S},s \Vdash \nabla\Phi$ if and only if, for all (some) successors σ of s$\in$S, $[\Phi,\sigma(s)\in\mathbf{E}(\Vdash_{\mathbb{A}})]$ (op. cit.).

A coalgebraic model of deterministic automata can finally be thus defined (391). An automaton is a tuple, $\mathbb{A} = \langle A, a_I, C, \delta, F\rangle$, such that A is the state space of the automaton $\mathbb{A}$; $a_I \in A$ is the automaton's initial state; C is the coding for the automaton's alphabet, mapping numerals to properties of the natural numbers; δ: A X C $\to$ A is a transition function, and F$\subseteq$ A is the collection of admissible states, where F maps A to $\{1,0\}$, such that F: A $\to$ 1 if a$\in$F and A $\to$ 0 if a$\notin$F (op. cit.).

The crux of the reconciliation between algebraic models of cognitivism and the formal foundations of modal expressivism is based on the dual isomorphism between categories of algebras and coalgebras: $\mathbb{A} = \langle A, \alpha{:}A \to \mathbf{E}(A)\rangle$ is dually isomorphic to the category of algebras over the functor α (417-418). For a category C, object A, and endofunctor $\mathbf{E}$, define a new arrow, α, s.t. $\alpha{:}\mathbf{E}A \to A$. A homomorphism, f, can further be defined between algebras $\langle A, \alpha\rangle$, and $\langle B, \beta\rangle$. Then, for the category of algebras, the following commutative square can be defined: (i) $\mathbf{E}A \to \mathbf{E}B$ ($\mathbf{E}f$); (ii) $\mathbf{E}A \to A$ (α); (iii) $\mathbf{E}B \to B$ (β); and (iv) $A \to B$ (f) (cf. Hughes, 2001: 7-8). The same commutative square holds for the category of coalgebras, such that the latter are defined by inverting the direction of the morphisms in both (ii) $[A \to \mathbf{E}A$ (α)], and (iii) $[B \to \mathbf{E}B$ (β)] (op. cit.).[6]

The significance of the foregoing is twofold. First and foremost, the above demonstrates how a formal correspondence can be effected between algebraic models of cognition and coalgebraic models which provide a natural setting for modal logics and automata. The second aspect of the philosophical significance of modal colagebraic automata is that – as a model of modal expressivism – the proposal is able to countenance fundamental properties in the foundations of mathematics, and circumvent the issues accruing to the attempt so to do by the competing expressivist approaches.

3.2   Material Adequacy

The material adequacy of epistemic modal algebras as a fragment of the representational theory of mind is witnessed by the prevalence of possible worlds semantics – the model theory for which is algebraic (cf. Blackburn et al., 2001: ch. 5) – in cognitive psychology, and the convergences thereof with theoretical physics.

In Bayesian perceptual psychology, e.g., the visual system is presented with a prior distribution of possibilities concerning the direction of a source of light. The set of possibilities is pointed, as the visual system calculates the likelihood that one of the possibilities is actual, and places a condition thereby on the accuracy of the attribution of properties – such as boundedness and

---

[6]Note that the theory for the model will be a topological coalgebra – cf. Takeuchi, 1985 for further discussion – and that, unlike isomorphism, the property of being dually isomorphic is not one-one (cf. Hughes, op. cit.: 13). Dual isomorphisms are thus more akin to bisimulations (cf. Blackburn et al., 2001: 64-66).

volume – to distal particulars (cf. Mamassian et al., 2000).

In the philosophy of physics, modal notions play, further, an ineliminable role in physical ontology and in the theoretical and observational characterizations thereof. Transition functions in quantum information theory can, for example, be interpreted counterfactually, in order to countenance the possible transformations of systems, the inputs of which include the expectation values of spin-state vectors (Deutsch, 2013). In the interpretation of physical theories, Ruetsche (2011: 9) argues, further, that '[T]o interpret a physical theory is to characterize the worlds possible according to that theory'. She notes that the interpretation of a physical theory falls into two phases, both of which constitutively involve modal notions. The first phase consists in the specification of the possible states of the models at issue; the possible values taken with regard to the observable kinematics thereof, e.g., the expectation values of Hilbert operators; and then the laws comprising the dynamic transformations therein (op. cit.). The second phase targets the nomologically possible worlds in which the properties of the first phase of interpretation are satisfied (op. cit.). In quantum field theory, unital algebras are availed of, as well, where the algebras are closed under commutativity and associativity; scalar multiplication by complex numbers, such that for all vectors A,B, and complex numbers c_n: $c_1(A + B) = c_1A + c_1B$; $(c_1 + c_2)A = c_1A + c_2A$; $c_1(c_2A) = (c_1c_2)A$; $(c_1A)B - A(c_1B) = c_1(AB)$; and multiplicative identity, i.e., for an element I, $AI = IA = A$ (74). Finally, Belot (2011) argues that – in the dispute between relationalist and substantivalist conceptions of spatial geometry – a relationalist proposal can be proffered, according to which – by analogy with nomological possibilities – the different metric relations that can be grounded in the intrinsic points comprising spatial regions are geometric possibilities.

Finally, in artificial intelligence, the subfield of knowledge representation draws on epistemic logic, where possibility and necessity are interpreted as belief and knowledge (Meyer and van der Hoeck, 1995; Fagin et al., 1995). Possibility and necessity may receive other interpretations in mental terms, such as that of conceivability and apriority (i.e. truth in all epistemic possibilities, or inconceivability that not ϕ). The Language of Thought hypothesis maintains that thinking occurs in a mental language with a computational syntax and a semantics. The philosophical significance of cognitivism about epistemic modality is that it construes epistemic intensions as abstract, computational functions in the mind, and thus provides an explanation of the relation that human beings bear to epistemic possibilities, rather than leaving

questions of the ground of epistemic possibilities unexplained. Cognitivism about epistemic modality argues that thoughts are comprised of intensions, or functions from epistemic possibilities to extensions. Cognitivism about epistemic modality provides a metaphysical explanation or account of the ground of thoughts, arguing that they are grounded in epistemic possibilities and intensions, which are themselves internal representations comprising the semantics for a mental language. This is consistent with belief and knowledge being countenanced in an epistemic logic for artificial intelligence. Epistemic possibilities are constitutively related thoughts, and figure furthermore in the analysis of notions such as apriority and conceivability, as well as belief and knowledge in epistemic logic for artificial intelligence.

In the the remainder of the paper, I outline an expressivist semantics for epistemic modals. I endeavor, then, to demonstrate the advantages accruing to the present approach to countenancing modal expressivism via modal colagebraic automata, via a comparison of the theoretical strength of the proposal when applied to characterizing the fundamental properties of the foundations of mathematics, by contrast to the competing approaches to modal expressivism and the limits of their applications thereto.

3.2.1   Expressivist Semantics for Epistemic Modals

Let expressivism about a domain of discourse be the claim that an utterance from that domain expresses a mental state, rather than states a fact (Hawke and Steinert-Threlkeld, 2020). Hawke and Steinert-Threlkeld (op. cit.) distinguish between semantic expressivism and pragmatic expressivism. Expressivism about epistemic modality takes the property expressed by $\diamond\phi$ to be $\{\mathbf{s}\subseteq W\colon \mathbf{s}\nVdash \neg\mathrm{p}\}$, where $\mathbf{s}$ is a state of information, W is a set of possible worlds, and $\mathbf{s}\Vdash\phi$ if and only if ϕ is assertible relative to $\mathbf{s}$, if and only if the state of information is compatible with ϕ. Semantic expressivism incorporates a "psychologistic semantics" according to which the value of ϕ is a partial function from information states to truth-values (op. cit.), such that "the mental type expressed by ϕ is characterized in terms of the assertibility relation $\Vdash$" and "the definition of $\Vdash$ is an essential part of that of $[\![\ ]\!]$" (op. cit.). Pragmatic expressivism rejects the psychologistic semantics condition, and "allows for a *gap* between the compositional semantic theory and $\Vdash$" (op. cit.).

Hawke and Steinert-Threlkeld (op. cit.) argue that satisfying the following conditions is a desideratum of any expressivist account about epistemic

possibility:

(Weak) Wide-scope Free Choice (**WFC**):
$\diamond p \vee \diamond \neg p \Vdash \diamond p \wedge \diamond \neg p$
Disjunctive Inheritence (**DIN**):
$(\diamond p \wedge q) \vee r \Vdash [\diamond(p \wedge q) \wedge q] \vee r$
Disjunctive Syllogism and Schroeder's Constraints:
DSF $\{\diamond \neg q, p \vee \Box q \nVdash p\}$
SCH $\{\diamond \neg p, p \vee \Box q \nVdash \Box q\}$

DSF and **SCH** record the failure of disjunctive syllogism in the presence of epistemic contradictions.

WFC is vindicated by the contention that when someone asserts $p \vee \neg p$, they neither believe p nor believe $\neg p$, and so are in a position to assert both $\diamond p$ and $\diamond \neg p$.

DIN is vindicated by the equivalence of the content of the utterances, e.g.,

(1) Nataly is at home and might be watching a film.

(2) Nataly is at home and might be watching a film at home.

Hawke and Steinert-Threlkeld's modal propositional assertibility semantics is then as follows.

Reading $t \subseteq s$: $[\![\phi]\!]^t \neq 1$ as "s refutes ϕ":

- if p is an atom: $[\![p]\!]^s = 1$ iff $s \subseteq V(p)$

  if p is an atom $[\![p]\!]^s = 0$ iff s refutes p

- $[\![\neg \phi]\!]^s = 1$ iff $[\![\phi]\!]^s = 0$

  $[\![\neg \phi]\!]^s = 0$ iff $[\![\phi]\!]^s = 1$

- $[\![\phi \wedge \psi]\!]^s = 1$ iff $[\![\phi]\!]^s = 1$ and $[\![\psi]\!]^s = 1$

  $[\![\phi \wedge \psi]\!]^s = 0$ iff s refutes $\phi \wedge \psi$

- $[\![\phi \vee \psi]\!]^s = 1$ iff there exists s_1, s_2 such that $s = s_1 \cup s_2$, $[\![\phi]\!]^{s_1} = 1$ and $[\![\psi]\!]^{s_2} = 1$

  $[\![\phi \vee \psi]\!]^s = 0$ iff s refutes $\phi \vee \psi$

- $[\![\diamond \phi]\!]^s = 1$ iff $[\![\phi]\!]^s \neq 0$

  $[\![\diamond \phi]\!]^s = 0$ iff s refutes $\diamond \phi$

- $\Box\phi := \neg\Diamond\neg\phi$

Unlike Yalcin's (2007) domain semantics, Veltman's (1996) update semantics, and Moss' (2015, 2018) probabilistic semantic expressivism, Hawke and Steinert-Threlkeld's assertibility semantics satisfies **WFC**, **DIN**, **DSF**, and **SCH** (op. cit.). As a preliminary, suppose

Proposition 1 If ϕ is $\Diamond$-free, then s $\Vdash \Diamond\phi$ holds iff there exists w$\in$s such that: {w} $\Vdash \phi$ (op. cit.).

Proof: s $\Vdash \Diamond\phi$ holds iff $[\![\phi]\!]^s \neq 0$. $[\![\phi]\!]^s = 0$ iff $[\![\phi]\!]^{\{w\}} = 0$ for every w$\in$s. So, $[\![\phi]\!]^s \neq 0$ iff $[\![\phi]\!]^w \neq 0$ for some w$\in$s iff {w} $\Vdash \phi$ for some w$\in$s (op. cit.).

For **WFC**, suppose that s $\Vdash \Diamond$p $\vee$ $\Diamond\neg$p. So, there exists s_1, s_2 that cover s and $s_1 \Vdash \Diamond$p and $s_2 \Vdash \Diamond\neg$p. By Proposition 1, there exist u,v$\in$s such that {u} $\Vdash$ p and {v} $\Vdash \neg$p. Thus, s $\Vdash \Diamond$p and s $\Vdash \Diamond\neg$p (op. cit.).

For **DIN**, suppose that s $\Vdash$ ($\Diamond$p $\wedge$ q) $\vee$ r. So, there exists s_1, s_2, such that s = $s_1 \cup s_2$ with $s_1 \Vdash \Diamond$p, $s_1 \Vdash$ q, and $s_2 \Vdash$ r. For every w$\in s_1$, {w} $\Vdash$ q. There also exists u$\in s_1$ such that {u} $\Vdash$ p. Hence, {u} $\Vdash$ p $\wedge$ q and – by Proposition 1 – $s_1 \Vdash \Diamond$(p $\wedge$ q). Thus s $\Vdash$ [$\Diamond$(p $\wedge$ q) $\wedge$ q] $\vee$ r (op. cit.).

For **DSF** and **SCH**, suppose that there is an s such that every world in s is either a p $\wedge$ $\neg$q world or a $\neg$p $\wedge$ q world. Suppose that there exists at least one p $\wedge$ $\neg$q world in s and at least one $\neg$p $\wedge$ q world in s (op. cit.).

Relativists about epistemic modals either relativize content or relativize truth to a context of assessment (Starr, 2012: 3; Egan and Weatherson, 2011: 11-14). According to content relativism, epistemic modals express different propositions in different contexts of assessment (Starr, op. cit.). According to truth relativism, epistemic modals express the same proposition, which is true relative to some assessors and false relative to others, such that truth is a three-place relation between a world, a judge, and a proposition, i.e. a centered world and a proposition (Starr, op. cit.: 3, 5). Thus, X believes that stealing is wrong is an ascription of belief in a centered proposition, i.e. a de se belief (Beddor, forthcoming: 9). As Yalcin (2011: 307) points out, utterances with epistemic modals on the truth relativist proposal thus express second-order states, the ascription of a de se belief (cf. Beddor, op. cit.).

That epistemic modal beliefs are second-order on the truth relativist proposal adduces against the merits of the view. Yalcin (op. cit.: 308) argues that non-human animals can entertain states expressed by epistemic modals,

and we here follow him in thinking that, by taking epistemic modal beliefs to be second-order de se ascriptions, the truth relativist proposal would preclude young children and non-human animals from entertaining epistemic possibilities. However, young children and non-human animals, while lacking the capacity to entertain second-order states, nevertheless entertain epistemic possibilities. The foregoing thus adduces in favor of the expressivist proposal that epistemic modals express first-order states of mind.

3.3   Modal Expressivism and the Philosophy of Mathematics

When modal expressivism is modeled via speech acts on a common ground of presuppositions, the application thereof to the foundations of mathematics is limited both (i) by the manner in which necessary propositions are characterized, as well as (ii) by the metalinguistic – rather than, e.g., epistemic – interpretation of the semantics.[7]

Because for example a proposition is taken, according to the proposal, to be identical to a set of possible worlds, all necessarily true mathematical formulas can only express a single proposition; namely, the set of all possible worlds (cf. Stalnaker, 1978; 2003: 51). Thus, although distinct set-forming operations will be codified by distinct axioms of a language of set theory, the axioms will be assumed to express the same proposition: The axiom of Pairing in set theory – which states that a unique set can be formed by combining an element from each of two extant sets: $\exists x \forall u(u \in x \iff u = a \lor u = b)$ – will be supposed to express the same proposition as the Power Set axiom – which states that a set can be formed by taking the set of all subsets

---

[7]See Chapter **11** for an application of the epistemic interpretation of two-dimensional intensional semantics to account for the modal profile of Orey sentences; i.e. mathematical propositions that are undecidable relative to the axioms of a given language. (For the origins of two-dimensional intensional semantics, see Kamp, 1967; Vlach, 1973; and Segerberg, 1973.) The distinction between epistemic and metaphysical possibilities, as they pertain to the values of mathematical formulas, is anticipated by Gödel's (1951: 11-12) distinction between mathematics in its subjective and objective senses, where the former concerns decidable formulas, and the latter records the values of formulas defined, owing to the incompleteness theorems, in a variant of the language augmented by stronger axioms of infinity. Axioms of infinity take the form, '$\exists x \varnothing \in x \land \forall u(u \in x \to \{u\} \in x)$', and record closure on the singleton of a real in the cumulative hierarchy of sets, where the real can be a large cardinal.

of an extant set: $\exists x \forall u(u \in x \iff u \subseteq a)$. However, that distinct operations – i.e., the formation of a set by selecting elements from two extant sets, by contrast to forming a set by collecting all of the subsets of a single extant set – are characterized by the different axioms is readily apparent.

Stalnaker endeavors to redress the objection by availing of the metasemantic interpretation of two-dimensional intensional semantics, according to which the necessarily true propositions comprising the common ground are yet consistent with contingently valued speech acts – e.g., assertions – about those propositions (2003: 54). Although it does record a difference in the epistemic status of agents with regard to whether they know or believe the propositions at issue, the proposal is yet not itself sufficient to distinguish between the senses of distinct, albeit necessarily true formulas.

Further, the metalinguistic interpretation of the semantics places a restriction on the expressive capacity of the mathematical languages at issue. By contrast to the objectual interpretation of the quantifiers, the metalinguistic interpretation of two-dimensional intensional semantics – by restricting itself to the remit of natural language semantics – is limited to substitutional quantification; i.e. to quantification over only those objects for which a name, in the natural language, has been specified (op. cit.). Given, however, the above restriction, substitutional quantification cannot account for impredicative comprehension; i.e., for the specification of extensions of terms, with reference to the totality of items of the type that the terms are intended to designate.

By contrast to the limits of Stalnaker's approach to modal expressivism, differences in the senses of formulas can, as noted, be recorded by the correspondence, in modal algebras, between modal operators and intensional functions, such that the senses of materially equivalent propositions-as-possibilities can shift when the possibilities are translated as intensions and figure within hyperintensional contexts. Further, both epistemic modal algebras and their dually corresponding modal coalgebraic automata permit of unrestricted objectual quantification, and are thus unconfined to the expressive limitations of the substitutional quantifiers of natural language.

A variation of Stalnaker's semantics is proffered in Yalcin (2016). Yalcin (op. cit.) argues that concepts and beliefs can be modeled in the manner of subject matters (cf. Lewis, 1998; Yablo, 2014), where the latter are interrogative updates on a background set of epistemically possible worlds, and the inquiry whether ϕ induces a focus on a subset of worlds which answer whether ϕ. In the setting of philosophy of mathematics, however, it is un-

clear how identifying subject matters with fragments of an epistemic modal space induced by interrogative updates can model the nature of, inter alia, Orey sentences such as the generalized continuum hypothesis (i.e., $\aleph_\alpha{}^{\aleph_\alpha} = \aleph_\alpha + 1$),

whose values are indeterminate relative to the present axioms of ZF set theory with choice, and the reduction in the incompleteness of which depends upon augmenting the relevant mathematical languages with stronger axioms of infinity (cf. Woodin, 2010). In order to determine the truth-values of Orey sentences, the space of epistemic modality might require an expanding domain by way of which the new axioms and their corresponding theorems can be accommodated. It is thus unclear how a treatment of subject matters as an induced *restriction* on the space of epistemic possibilities – i.e., as interrogative-induced operations from a set of possibilities to any particular subset thereof – could begin to account for the required expansion in epistemic states.

Thomasson (2007) argues for a version of modal expressivism which she refers to as 'modal normativism', according to which alethic modalities are to be replaced by deontic modalities taking the form of object-language, modal indicative conditionals (op. cit.: 136, 138, 141). The modal indicative conditionals serve to express constitutive rules pertaining, e.g., to ontological dependencies which state that: 'Necessarily, if an entity satisfying a property exists then a distinct entity satisfying a property exists' (143-144), and generalizes to other expressions, such as analytic conditionals which state, e.g., that: 'Necessarily, if an entity satisfies a property, such as being a bachelor, then the entity satisfies a distinct yet co-extensive property, such as being unmarried' (148). A virtue of Thomasson's interpretation of modal indicative conditionals as expressing both analytic and ontological dependencies is that it would appear to converge with the 'If-thenist' proposal in the philosophy of mathematics. 'If-thenism' is an approach according to which, if an axiomatized mathematical language is consistent, then (i) one can either bear epistemic attitudes, such as fictive acceptance, toward the target system (cf. Leng, 2010: 180) or (ii) the system (possibly) exists [cf. Russell (op. cit.: §1); Hilbert (1899/1980: 39); Menger (1930/1979: 57); Putnam (1967); Shapiro (2000: 95); Chihara (2004: Ch. 10); and Awodey (2004: 60-61)].[8] However, there are at least two issues for the modal normativist approach

---

[8]See Leng (2009), for further discussion. Field (1980/2016: 11-21; 1989: 54-65, 240-241) argues in favor of the stronger notion of conservativeness, according to which consistent mathematical theories must be satisfiable by internally consistent theories of physics. More

in the setting of the philosophy of mathematics. One general issue for the proposal is that the treatment of quantification remains unaddressed, given that there are translations from modal operators, such as figure in modal indicatives, into existential and universal quantifiers.[9] A second issue for the normative indicative conditional approach is that Thomasson's normative modalities are unimodal. They are thus not sufficiently fine-grained to capture distinctions such as Gödel's (op. cit.) between mathematics in its subjective and objective senses.[10] Further distinctions between the types of mathematical modality can be delineated which permit epistemic types of mathematical possibility to serve as a guide as to whether a formula is metaphysically mathematically possible.[11] The convergence between epistemic and metaphysical mathematical modalities can be countenanced via a two-dimensional intensional semantics. Thus, by eschewing alethic modalities for unimodal, normative indicatives, the normative modalities are unable to account for the relation between the alethic interpretation of modality and, e.g., logical mathematical modalities treated as consistency operators on languages (cf. Field, 1989: 249-250, 257-260; Leng: 2007; 2010: 258), or for the

generally, for a class of assertions, A, comprising a theory of fundamental physics, and a class of sentences comprising a mathematical language, M, any sentences derivable from A+M ought to be derivable from A alone. Another variation on the 'If-thenist' proposal is witnessed in Field (2001: 333-338), who argues that the existence of consistent forcing extensions of set-theoretic ground models adduces in favor of there being a set-theoretic pluriverse, and thus entrains indeterminacy in the truth-values of undecidable sentences. For a similar proposal, which emphasizes the epistemic role of examining how instances of undecidable sentences obtain and fail so to do relative to forcing extensions in the set-theoretic pluriverse, see Hamkins (2012: §7).

[9]The formal correspondence between modalities and quantifiers is anticipated by Aristotle (*De Interpretatione*, 9; *De Caelo*, I.12), who defines the metaphysical necessity of a proposition as its being true at all times. For detailed discussion of Aristotle's theory, see Waterlow (1982). For a contemporary account of the multi-modal logic for metaphysical and temporal modalities, see Dorr and Goodman (ms). For contemporary accounts of the correspondence between modal operators and quantifiers see von Wright (1952/1957), where von Wright anticipates Kripke's (1963) relational semantics for modal logic; Montague (1960/1974: 75); Lewis (1975/1998; 1981/1998); Kratzer (op. cit.; 1981/2012); and Kuhn (1980).

[10]See footnote 7.

[11]See Author (ms), for further discussion. A precedent is Reinhardt (1974: 199-200), who proposes the use of imaginary sets, classes, and projections, as 'imaginary experiments' (204), in order to ascertain the consequences of accepting new axioms for ZF which might account for the reduction of the incompleteness of Orey sentences. See Maddy (1988,b), for critical discussion.

convergence between epistemic possibilities concerning decidability and their bearing on the metaphysical modal status of undecidable sentences.

According, finally, to Brandom's (op. cit.) modal expressivist approach, terms are individuated by their rules of inference, where the rules are taken to have a modal profile translatable into the counterfactual forms taken by the transition functions of automata (cf. Brandom, 2008: 142). In order to countenance the metasemantic truth-conditions for the object-level, pragmatic abilities captured by the automata's counterfactual transition states, Brandom augments a first-order language comprised of a stock of atomic formulas with an incompatibility function (141). An incompatibility function, I, is defined as the incoherence of the union of two sentences, where incoherence is a generalization of the notion of inconsistency to nonlogical vocabulary.

x $\cup$ y $\in Inc \iff$ x $\in I$(y) (141-142).

Incompatibility is supposed to be a modal notion, such that the union of the two sentences is incompossible (126). A sentence, β is an incompatibility-consequence, $\Vdash_I$, of a sentence, α, iff there is no sequence of sentences, $<\gamma_1$, ..., $\gamma_n>$, such that it can be the case that $\alpha \Vdash_I <\gamma_1$, ..., $\gamma_n>$, yet not be the case that $\beta \Vdash_I <\gamma_1$, ..., $\gamma_n>$ (125). To be incompatible with a necessary formula is to be compatible with everything that does not entail the formula (129-130). Dually, to be incompatible with a possible formula is to be incompatible with everything compatible with something compatible with the formula (op. cit.).

There are at least two, general issues for the application of Brandom's modal expressivism to the foundations of mathematics.

The first issue is that the mathematical vocabulary – e.g., the set-membership relation, $\in$ – is axiomatically defined. I.e., the membership relation is defined by, inter alia, the Pairing and Power Set axioms of set-theoretic languages. Thus, mathematical terms have their extensions individuated by the axioms of the language, rather than via a set of inference rules that can be specified in the absence of the mention of truth values. Even, furthermore, if one were to avail of modal notions in order to countenance the intensions of the mathematical vocabulary at issue – i.e., functions from terms in intensional contexts to their extensions – the modal profile of the intensions is orthogonal to the properties encoded by the incompatibility function. Fine (2006) avails, e.g., of a dynamic logic in order to countenance the possibility of reinterpreting the intensions at issue, and of thus accounting for variance in the range of the domains of quantifier expressions. The dynamic possibilities are

specified as operational conditions on tracking increases in the size of the cardinality of the universe (Fine, 2005). Uzquiano (2015,a) argues that it is always possible to reinterpret the intensions of non-logical vocabulary, as one augments one's language with stronger axioms of infinity and climbs thereby farther up the cumulative hierarchy of sets. The reinterpretations of, e.g., the concept of set are effected by the addition of new large cardinal axioms, which stipulate the existence of larger inaccessible cardinals. However, it is unclear how the incompatibility function – i.e., a modal operator defined via Boolean negation and a generalized condition on inconsistency – might similarly be able to model the intensions pertaining to the ontological expansion of the cumulative hierarchy.

The second issue is that the truth-conditional formulas in the metalanguage of Brandom's modal expressivist semantics are not compositional. I.e., it is not the case of the clauses in the metalanguage that their truth-conditions are formed by the composition of the semantic values of their component expressions. While the formulas are recursively formed – because the decomposition of complex formulas into atomic formulas is decidable although computationally infeasible – formulas in the language are not compositional, because they fail to satisfy the subformula property to the effect that the value of a logically complex formula is calculated as a function of the values of the component logical connectives applied to subformulas therein (135).[12]

By contrast to the limits of Brandom's approach to modal expressivism, modal coalgebraic automata can circumvent both of the issues mentioned in the foregoing. In response to the first issue, concerning the axiomatic individuation and intensional profiles of mathematical terms, functors of modal coalgebraic automata can be interpreted in order to provide a precise delineation of the intensions of the target vocabulary (cf. Chapter **11**). In response, finally, to the second of the above issues, the values taken by modal coalgebraic automata are both decidable and computationally feasible, while the dual isomorphism of colagebras to Boolean-valued models of modal algebras ensures that the formulas therein retain their compositionality. The decidability of colagebraic automata can further be witnessed by the role of modal coalgebras in countenancing the modal profile of Ω-logical conse-

---

[12]Let a decision problem be a propositional function which is feasibly decidable, if it is a member of the polynomial time complexity class; i.e., if it can be calculated as a polynomial function of the size of the formula's input (cf. Dean, 2015).

quence, where – given a proper class of Woodin cardinals – the values of mathematical formulas can remain invariant throughout extensions of the ground models comprising the set-theoretic pluriverse (cf. Woodin, 2010; Chapter **11**). The individuation of large cardinals can further be characterized by the functors of modal coalgebraic automata, when the latter are interpreted so as to countenance the elementary embeddings constitutive of large cardinal axioms in the category of sets.

3.4   Concluding Remarks

In this essay, I have endeavored to account for a mathematically tractable background against which to model both modal cognitivism and modal expressivism. I availed, to that end, of the dual isomorphism between epistemic modal algebras and modal coalgebraic automata. Epistemic modal algebras were shown to comprise a materially adequate fragment of the language of thought, given that models thereof figure in both the cognitive, as well as in the interpretations of the physical, sciences. It was then shown how the approach to modal expressivism here proffered, as regimented by the modal coalgebraic automata to which the epistemic modal algebras are isomorphic, avoids the pitfalls attending the competing modal expressivist approaches based upon both the inferentialist approach to concept-individuation and the approach to codifying the speech acts in natural language via intensional semantics. The present modal expressivist approach was shown, e.g., to avoid the limits of the foregoing in the philosophy of language, as they concerned the status of necessary propositions; substitutional quantification; the in-applicability of inferentialist-individuation to mathematical vocabulary; and failures of compositionality. Countenancing modal expressivism via modal coalgebraic automata was shown, then, to be able to account for both the intensions of mathematical terms and possible reinterpretations thereof; for the modal profile of Ω-logical consequence in the category of sets; and for the elementary embeddings constitutive of large cardinal axioms in set-theoretic languages.|

Part II: Epistemic Modality and Conceivability

Chapter 4

Conceivability and Haecceitism

4.1 Introduction

In this essay, I endeavor to provide an account of how the epistemic interpretation of two-dimensional intensional semantics can be sensitive to *de re* modalities. Let a model, M, be comprised of a set of epistemically possible worlds C; a set of metaphysically possible worlds P; a domain, D, of terms and formulas; binary relations defined on each of C and P; and a valuation function mapping terms and formulas to subsets of C and P, respectively. So, $M = \langle C, P, D, R_C, R_P, V \rangle$. A term or formula is epistemically necessary or apriori iff it is inconceivable for it to be false ($\Box \iff \neg \Diamond \neg$). A term or formula is negatively conceivable iff nothing rules it out apriori ($\Diamond \iff \neg \Box \neg$). A term or formula is positively conceivable only if the term or formula can be perceptually imagined. According to the epistemic interpretation of two-dimensional intensional semantics, the semantic value of a term or formula can then be defined relative to two parameters, a context and an index. The context ranges over the set of epistemically possible worlds, and the index ranges over the set of metaphysically possible worlds. The value of the term or formula relative to the context determines the value of the term or formula relative to the index. Thus, the epistemically possible value of the term or formula constrains the metaphysically possible value of the term or formula; and so conceivability might, given the foregoing, serve as a guide to metaphysical possibility.

Roca-Royes (2011) and Chalmers (2010; 2011; 2014) note that, on the above semantics, epistemic possibility cannot track the difference between the

metaphysical modal profile of a non-essential proposition – e.g., that there is a shooting star – and the metaphysical modal profile of an essential definition, such as a theoretical identity statement – e.g., that water = H2O. Another principle of modal metaphysics to which epistemic possibilities are purported to be insensitive is haecceity comprehension; namely, that $\Box\forall x,y\Box\exists\Phi(\Phi x \iff x = y)$.

The aim of this note is to redress the contention that epistemic possibility cannot be a guide to the principles of modal metaphysics. I will argue that the interaction between the two-dimensional intensional framework and the mereological parthood relation enables epistemic possibilities to target the haecceitistic properties of individuals.

In Section **2**, I examine a necessary condition on admissible cases of conceivability entailing metaphysical possibility in the two-dimensional intensional framework, focusing on the property of super-rigidity. I argue that – despite the scarcity of properties which satisfy the super-rigidity condition – metaphysical properties such as the parthood relation do so. In Section **3**, I address objections to two dogmas of the semantic rationalism underpinning the epistemic interpretation of two-dimensional intensional semantics. The first dogma states that distinctions can be delineated between linguistic intensions and conceptual epistemic intensions, while the second dogma records that there are criteria on the basis of which formal from informal domains, unique to the extensions of various concepts, can be distinguished, such that the modal profiles of those concepts would thus be determinate. I examine the Julius Caesar problem as a test case. I specify, then, a two-dimensional intensional formula encoding the relation between the epistemic possibility of haecceity comprehension and its metaphysical possibility. In Section **4**, I address objections from the indeterminacy of ontological principles relative to the space of epistemic possibilities, and from the consistency of epistemic modal space. Section **5** provides concluding remarks.

4.2   Super-rigidity

The interaction between mereological parthood and epistemic and metaphysical possibility avoids one crucial issue for the epistemic interpretation of two-dimensional intensional semantics. The issue is that, unless the semanic value for a term is 'super-rigid', i.e. maps to the same extension throughout the sets of epistemic and metaphysical possibilities, the extension of the term

in epistemic modal space risks diverging from the extension of the term in metaphysical modal space.

There appear to be only a few expressions which satisfy the super-rigidity condition. Such terms include those referring to the properties of phenomenal consciousness, to the parthood relation, and perhaps to the property of friendship (Chalmers, 2012: 367, 374). Other candidates for super-rigidity are taken to include metaphysical terms such as 'cause' and 'fundamental'; numerical terms such as 'one'; and logical constants such as '$\wedge$' (Chalmers, op. cit.). However, there are objections to each of the other proposed candidates.

Against the super-rigidity of 'fundamental', Fine (2001: 3) argues that a proposition is fundamental if and only if it is real, while Sider (2011: 112, 118) argues that a proposition is fundamental iff it possesses a truth-condition (in a 'metaphysical semantics', stated in perfectly joint-carving terms) for the sub-propositional entities – expressed by quantifiers, functions, predicates – comprising the target proposition. The absolute joint-carving terms are taken to include logical vocabulary (including quantifiers), metaphysical predicates such as mereological parthood, and physical predicates.

Against the super-rigidity of 'cause', Sider (op.cit.: 8.3.5) notes that a causal deflationist might argue that causation is non-fundamental. By contrast, a causal nihilist might argue that causation is non-fundamental as well, though for the distinct reason that there is no causation. So, while both the deflationist and nihilist believe that 'cause' does not carve at the joints – the nihilist can still state that there is a related predicate, 'cause*', such that they can make the joint-carving claim that 'Nothing causes* anything', whereas the deflationist will remain silent, and maintain that no broadly causal locutions carve at the joints.

Against the super-rigidity of 'one', Benacerraf (1965) notes that, in the reduction of number theory to set theory, there must be, and is not, a principled reason for which to prefer the identification of natural numbers with von Neumann ordinals (e.g., $2 = \{\emptyset, \emptyset\}$), rather than with Zermelo ordinals (i.e., an order-type of a well-ordering $2 = \{\{\emptyset\}\}$).[1]

Against the super-rigidity of the logical connective, $\wedge$, the proponent of model-theoretic validity will prefer a definition of the constant according to which, for propositions ϕ and ψ and a model, M, M validates $\phi \wedge \psi$ iff M

---

[1]Cf. Zermelo (1908/1967) and von Neumann (1923/1967). Well-orderings are irreflexive, transitive, binary relations on all non-empty sets, defining a least element in the sets.

validates ϕ and M validates ψ. By contrast, the proponent of proof-theoretic validity will prefer a distinct definition which makes no reference to truth, according to which $\wedge$ is defined by its introduction and elimination rules: $\phi,\psi \vdash \phi\wedge\psi$; $\phi\wedge\psi \vdash \phi$; $\phi\wedge\psi \vdash \psi$.

Finally, terms for physical entities such as 'tensor field' might have a rigid intension mapping the term to the same extension in metaphysical modal space, and a non-rigid intension mapping the term to distinct extensions in epistemically possible space, such that what is known about the term is contingent and might diverge from its necessary metaphysical profile.[2] That physical terms are not super-rigid might be one way to challenge the soundness of the conceivability argument to the effect that, if it is epistemically possible that truths about consciousness cannot be derived from truths about physics, then the dissociation between phenomenal and physical truths is metaphysically possible (cf. Chalmers, 2010: 151).

Crucially for the purposes of this note, there appear to be no clear counterexamples to the claim that mereological parthood is super-rigid. If this is correct, then mereological parthood in the space of epistemic modality can serve as a guide to the status of mereological parthood in metaphysical modal space. The philosophical significance of the foregoing is that it belies the contention proffered by Roca-Royes (op. cit.) and Chalmers (op. cit.) concerning the limits of conceivability-based modal epistemology. The super-rigidity of the parthood relation ensures that the interaction between the conceivability of mereological parthood, which records the existence of haecceities, can serve as a guide to the metaphysical modal profile of haecceity comprehension.

4.3   Two Dogmas of Semantic Rationalism

The tenability of the foregoing depends upon whether objections to what might be understood as the two dogmas of semantic rationalism can be circumvented.[3]

---

[2]A 'tensor field' is a function from m '1-forms' at a spacetime point, p, and n vectors at p, to the real numbers. A 1-form is a function, ω, s.t. ω maps four vectors to the real numbers, and satisfies the condition that for vectors ≥ 2, μ, τ, and real numbers α and β: $\omega(\alpha\mu + \beta\tau) = \alpha\omega(\mu) + \beta\omega(\tau)$. Cf. Arntzenius (2012): 72.

[3]Thanks to xx for the objections.

4.3.1 The First Dogma

The first dogma of semantic rationalism mirrors Quine's (1951) contention that one dogma of the empiricist approach is the distinction that it records between analytic and synthetic claims. The analogous dogma in the semantic rationalist setting is that a distinction can be drawn between linguistic epistemic intensions – witnessed by differences in the cognitive significance of two sentences or terms which have the same extension, e.g., with x = 2, 'x^2' and '2x' – by contrast to conceptual epistemic intensions – e.g., those which denote the properties of phenomenal consciousness. The distinction coincides with two interpretations of two-dimensional intensional semantics. As noted, the epistemic interpretation of two-dimensional intensional semantics takes the value of a formula relative to a context ranging over epistemically possible worlds to determine the extension of the formula relative to an index ranging over metaphysically possible worlds (cf. Chalmers, op. cit.). According to the metasemantic interpretation, a sentence, such as that 'water = H$_2$0', is metaphysically necessary, whereas assertions made about metaphysically necessary sentences record the non-ideal epistemic states of agents and are thus contingent (cf. Stalnaker, 1978, 2004). The first dogma is thus to the effect that there are distinct sets of worlds – sets of non-linguistic conceptual possibilities and of linguistic presuppositions, respectively – over which the context ranges in the epistemic and metasemantic interpretations.

Two examples might be apposite. The physical law that force can be identified by calculating the product of mass and acceleration, f = ma, has a distinct linguistic intension than that for a reformulated version of the law, according to which force can be identified by calculating the product of mass and the independently calculated product of acceleration and the second derivative of position, f = m(d^{2}x/dt^2) (cf. Hicks and Schaffer, 2015: 17). However, 'f = ma' and 'f = m(d^{2}x/dt^2)' have identical ideal conceptual epistemic intensions. Similarly, the linguistic intension for the parthood relation can vary while its conceptual intension remains constant. Thus, the linguistic intension for the sentence that 'the class of renates is a part of the class of cordates' is distinct from the linguistic intension for the sentence that 'the class of entities with kidneys is a part of the class of entities with hearts'. However, the ideal conceptual epistemic intensions for the two thoughts are identical.

If no conditions on the distinctness between linguistic and conceptual epistemic intensions can be provided, then variance in linguistic intension

might adduce against the uniqueness of the conceptual intension. Because of the possible proliferation of epistemic intensions, conditions on the superrigidity of the formulas and terms at issue might thereby not be satisfiable. The significance of the first dogma of semantic rationalism is that it guards against the collapse of conceptual and linguistic epistemic intensions, and thus the collapse of language and thought.

A defense of the first dogma of semantic rationalism might, in response, be proffered, in light of the status of higher-order distributive plural quantification in natural language semantics. Plural quantifiers are distributive, if the individuals comprising the plurality over which the quantifier ranges are conceived of singly, rather than interpreting the quantifier such that it ranges over irreducible collections. Natural language semantics permits plural quantification into both first and second-level predicate position. For all second-order variables ranging over a domain of individuals, one can define a predicate denoting a plurality of individuals at the first level (cf. Rayo, 2006). Subsequently, if there is an individual which satisfies the predicate, then the individual is a part of the extension of the predicate, i.e., the relevant plurality of individuals . For an example in English, the predicate can be interpreted such that it denotes the plurality, 'the books'. Then, an individual satisfies the predicate if and only if it is a part of the plurality of books. It is similarly innocuous in English to avail of plural quantification into second-level predicate position. For all third-order variables ranging over a domain of individuals, one can define a predicate denoting a plurality of individuals at the second level. E.g., the predicate might be interpreted such that it denotes the second-level plurality, the chamber pieces, in the sentence, 'These quartets are among her chamber pieces'. Then, the first-level plurality, the quartets, satisfies the predicate if and only if it is a part of the second-level plurality of chamber pieces over which the third-order plural quantifier has been defined to range (op. cit.).

Advancing to a higher order, for all fourth-order variables ranging over a domain of individuals, one can define a predicate which denotes a plurality of individuals at the third level. However, there are no examples of plural quantification into third-level predicate position in empirical linguistics, despite that examples thereof can be readily countenanced in intended models of formal languages.

The philosophical significance of the foregoing is that it is conceptually innocuous for plural terms of an arbitrary type, i_α, to satisfy a polyadic predicate at $i_{\alpha+2}$th-order which denotes pluralities of type $i_{\alpha+1}$, such that plural-

ities of the i_αth-type can be defined as parts of pluralities of the $i_{\alpha+1}$th-type. Yet – for polyadic predicates at order > 3 – the above has, as just noted, no analogue in natural language. As follows, higher-order plural quantification might adduce in favor of the first dogma of semantic rationalism, to the effect that linguistic and conceptual epistemic intensions can be sufficiently distinguished.

4.3.2   The Second Dogma

The second dogma of semantic rationalism mirrors Quine's (op. cit.) contention that another dogma of empiricism is the reduction of the meaning of a sentence to the empirical data which verifies its component expressions. The analogous dogma in the semantic rationalist setting states that individuation-conditions on concepts can be provided in order to distinguish between concepts unique to formal and informal domains. The significance of the second dogma of semantic rationalism is that whether the objects falling under a concept belong to a formal domain of inquiry will subsequently constrain its modal profile.

In the space of epistemic possibility, it is unclear, e.g., what reasons there might be to preclude implicit definitions such as that the real number of the x's is identical to Julius Caesar (cf. Frege, 1884/1980: 56; Clark, 2007) by contrast to being identical to a unique set of rational numbers as induced via Dedekind cuts. It is similarly unclear how to distinguish, in the space of epistemic possibility, between formal and informal concepts, in order to provide a principled account of when a concept, such as the concept of 'set', can be defined via the axioms of the language in which it figures, by contrast to concepts such as 'water', where definitions for the latter might target the observational, i.e. descriptive and functional, properties thereof.

The concept of mereological parthood provides a further borderline case. While the parthood relation can be axiomatized so as to reflect whether it is irreflexive, non-symmetric, and transitive, its status as a formal property is more elusive. The fact, e.g., that an order-type is part of the sequence of ordinal numbers impresses as being necessary, while yet the fact that a number of musicians comprise the parts of a chamber ensemble might impress as being contingent.

The Julius Caesar Problem

The Julius Caesar problem, and the subsequent issue of whether there might be criteria for delineating formal from informal concepts in the space of epistemic modality, may receive a unified response. The ambiguity with regard to whether the parthood relation is formal – given that its relata can include both formal and informal objects – is similar to the ambiguity pertaining to the nature of real numbers. As Frege (1893/2013: 161) notes: 'Instead of asking which properties an object must have in order to be a magnitude, one needs to ask: how must a concept be constituted in order for its extension to be a domain of magnitudes [. . .] a thing is a magnitude not in itself but only insofar as it belongs, with other objects, to a class that is a domain of magnitudes'. Frege defines a magnitude as the extension of a relation on arbitrary domains (op. cit.). The concept of a magnitude is then referred to as a 'Relation', and domains of magnitudes are defined as classes of Relations (162). Bypassing the rational numbers, Frege defines, then, the real numbers as relations on – namely, ratios of – magnitudes; and thus refers to the real numbers as 'Relations on Relations', because the extension of the higher-order concept of real number is taken to encompass the extension of the lower-order concept of classes of Relations, i.e., domains of magnitudes (op. cit.). The interest of Frege's definition of the concept of real number is that explicit mention must be made therein to a domain of concrete entities to which the number is supposed, as a type of measurement, to be applied.

In response: The following implicit definitions – i.e., abstraction principles – can be provided for the concept of real number, where the real numbers are defined as sets, or Dedekind cuts, of rational numbers. Following Shapiro (2000), let F,G, and R denote rational numbers, such that concepts of the reals can be specified as follows: $\forall$F,G[$\mathbf{C}$(F) = $\mathbf{C}$(G) $\iff$ $\forall$R(F$\leq$R $\iff$ G$\leq$R)]. Concepts of rational numbers can themselves be obtained via an abstraction principle in which they are identified with quotients of integers – [$\mathbf{Q}\langle$m,n$\rangle$ = $\mathbf{Q}\langle$p,q$\rangle$ $\iff$ n = 0 $\wedge$ q = 0 $\vee$ n $\neq$ 0 $\wedge$ q $\neq$ 0 $\wedge$ m x q = n x p]; concepts of the integers are obtained via an abstraction principle in which they are identified with differences of natural numbers – [$\mathbf{D}(\langle$x,y$\rangle$) = $\mathbf{D}(\langle$z,w$\rangle$) $\iff$ x + w = y + z]; concepts of the naturals are obtained via an abstraction principle in which they are identified with pairs of finite cardinals – '$\forall$x,y,z,w[$\langle$x,y$\rangle$(=$\mathbf{P}$) = $\langle$z,w$\rangle$(=$\mathbf{P}$) $\iff$ x = z $\wedge$ y = w]; and concepts of the cardinals are obtained via Hume's Principle, to the effect that cardinals are identical if and only if they are equinumerous – $\forall\mathbf{A}\forall\mathbf{B}\exists$R[[Nx: $\mathbf{A}$ = Nx:

B ≡ ∃R[∀x[**A**x → ∃y(**B**y ∧ Rxy ∧ ∀z(**B**z ∧ Rxz → y = z))] ∧ ∀y[**B**y → ∃x(**A**x ∧ Rxy ∧ ∀z(**A**z ∧ Rzy → x = z))]]].

Frege notes that 'we can never [...] decide by means of [implicit] definitions whether any concept has the number Julius Caesar belonging to it, or whether that same familiar conqueror of Gaul is a number or not' (1884/1980: 56). A programmatic line of response endeavors to redress the Julius Caesar problem by appealing to sortal concepts, where it is an essential property of objects that they fall in the extension of the concept (cf. Hale and Wright, 2001: 389, 395). In order further to develop the account, I propose to avail of recent work in which identity conditions are interpreted so as to reflect relations of essence and explanatory ground. The role of the essentiality operator will be to record a formal constraint on when an object falls under a concept 'in virtue of the nature of the object' (Fine, 1995: 241-242). The role of the grounding operator will be to record a condition on when two objects are the same, entraining a hyperintensional type of implicit definition for concepts which is thus finer-grained and less susceptible to error through misidentification.

In his (2015), Fine treats identity criteria as generic statements of ground. By contrast to **material** identity conditions which specify when two objects are identical, **criterial** identity conditions explain in virtue of what the two objects are the same. Arbitrary, or generic, objects are then argued to be constitutive of criterial identity conditions. Let a model, M, for a first-order language, L, be a tuple, where $M = \langle I, A, R, V \rangle$, with I a domain of concrete and abstract individuals, A a domain of arbitrary objects, R a dependence relation on arbitrary objects, and V a non-empty set of partial functions from A to I (cf. Fine, 1985). The arbitrary objects in A are reified variables. The dependence relation between any a and b in A can be interpreted as a relation of ontological dependence (op. cit.: 59-60). Informally, from $a \in A$ s.t. $F(a)$, one can infer $\forall x.F(x)$ and $\exists x.F(x)$, respectively (57). Then, given two arbitrary objects, x and y, with an individual i in their range, '[$(x = i \wedge y = i) \rightarrow x = y$]', such that x and y mapping to a common individual explains in virtue of what they are the same (Fine, 2015).

Abstraction principles for, e.g., the notion of set, as augmented so as to record distinctions pertaining to essence and ground, can then be specified as follows:

- Given x,y, with Set(x) ∧ Set(y): [∀z(z∈x ≡ z∈y) ←$_{x,y}$ (x = y)]

(Intuitively, where the 'given' expression is a quantifier ranging over the

domain of variables-as-arbitrary objects: Given x, y, whose values are sets, it is essential to x and y being the same that they share the same members); and

- Given x,y, with Set(x) ∧ Set(y): $[\forall z(z \in x \equiv z \in y) \rightarrow_{x,y} (x = y)]$

(Intuitively: Given arbitrary objects, x, y, whose values are sets, the fact that x and y share the same members grounds the fact that they are the same).

Combining both of the above directions yields the following hyperintensional, possibly asymmetric, biconditional:

- Given x,y, with Set(x) ∧ Set(y): $[\forall z(z \in x \equiv z \in y) \leftrightarrow_{x,y} (x = y)]$.

A reply to the Julius Caesar problem for real numbers might then avail of the foregoing metaphysical implicit definitions, such that the definition would record the essentiality to the reals of the property of being necessarily non-concrete:

- Given F,G$[\mathbf{C}(F) = \mathbf{C}(G) \leftrightarrow_{F,G} \forall R(F \leq R \iff G \leq R)]$, and

- $\Box \forall X_{X/F} \Box \exists Y [\neg E(Y) \wedge \Box(X = Y)]$

(Intuitively: Given arbitrary objects, F,G, whose values are the real numbers: It is essential to the F's and the G's that the concept of the Fs is identical to the concept of the G's iff (i) F and G are identical subsets of a limit rational number, R, and (ii) with E(x) a concreteness predicate, necessarily for all real numbers, X, necessarily there is a non-concrete object Y, to which necessarily X is identical; i.e., the reals are necessarily non-concrete. The foregoing is conversely the ground of the identification.)[4]

Heck (2011: 129) notes that the Caesar problem incorporates an epistemological objection: "Thus, one might think, there must be more to our apprehension of numbers than a mere recognition that they are the references of expressions governed by HP [Hume's Principle – HK]. Any complete

---

[4]Rosen and Yablo (2020) also avail of real, or essential, definitions in their attempt to solve the Caesar problem, although their real definitions do not target grounding-conditions. The need for a grounding-condition is mentioned in Wright (2020: 314, 318). The approach here developed, of solving the Caesar problem by availing of metaphysical definitions, was arrived at independently of Rosen and Yablo (op.cit.) and Wright (op. cit.).

account of our apprehension of numbers as objects must include an account of what distinguishes people from numbers. But HP alone yields no such explanation. That is why Frege writes: 'Naturally, no one is going to confuse [Caesar] with the [number zero]; but that is no thanks to our definition of [number]' (Gl, 62)".

The condition of being necessarily non-concrete in the metaphysical definition for real numbers – which includes conditions of essence and ground – provides a reply to the foregoing epistemological objection, i.e. the required account, beyond the abstraction principle, of what distinguishes people from numbers.

4.3.3   Mereological Parthood

The above proposal can then be generalized, in order to countenance the abstract profile of the mereological parthood relation. By augmenting the axioms for parthood in, e.g., classical mereological parthood with a clause to the effect that it is essential to the parthood relation that it is necessarily non-concrete, parthood can thus be understood to be abstract; and truths in which the relation figures would thereby be necessary.

- Given x: $\Phi(x) \wedge \Box\forall x\Box\exists y\ [\neg E(y) \wedge \Box(x = y)] \leftrightarrow_x \Gamma(x)$ where

- $\Gamma(x) :=$ x is the parthood relation, $<$, which is irreflexive, asymmetric, and transitive, and where the relation satisfies the axioms of classical extensional mereology codified by the predicate, $\Phi(x)$ (cf. Cotnoir, 2014):

  Weak Supplementation: $x < y \rightarrow \exists z[(z < y \vee z = y) \wedge \neg\exists w(w < z \vee w = z) \wedge (w < x \vee w = x)]$, and

  Unrestricted Fusion: $\forall xx\exists y[F(y,xx)]$,

  with the axiom of Fusion defined as follows:

  Fusion: $F(t,xx) := (xx < t \vee xx = t) \wedge \forall y[(y < t \vee y = t) \rightarrow (y < xx \vee y = xx)]$

Fusions are themselves abstracta, formed by a fusion-abstraction principle. The abstraction principle states that two singular terms – in which an abstraction operator, σ, from pluralities to fusions figures as a subformula –

are identical, if and only if the fusions overlap the same locations (cf. Cotnoir, ms). Let a topological model be a tuple, comprised of a set of points in topological space, μ; a domain of individuals, D; an accessibility relation, R; and a valuation function, V, assigning distributive pluralities of individuals in D to subsets of μ:

M = $\langle \mu, D, R, V \rangle$;

R = R(xx,yy)$_{xx,yy \in \mu}$ iff R$_{xx} \subseteq \mu_{xx}$ x μ_{xx}, s.t. if R(xx,yy), then $\exists o \subseteq \mu$, with xx$\in$o s.t. $\forall yy \in o R(xx,yy)$, where the set of points accessible from a privileged node in the space is said to be open; and V = $f(ii \in D, m \in \mu)$.[5] Necessity is interpreted as an interiority operator on the space:

M,xx $\Vdash \Box \phi$ iff $\exists o \subseteq \mu$, with xx$\in$o, such that $\forall yy \in o$ M,yy $\Vdash \phi$.

The following fusion abstraction principle can then be specified:

Given xx,yy,F[σ(xx,F) = σ(yy,F) $\leftrightarrow_{xx,yy}$ [f(xx,m_1) $\cap$ f(yy,m_1) ($\neq \emptyset$)]].

(Intuitively, given arbitrary objects whose values are the pluralities, xx,yy: It is essential to xx and yy that fusion-abstracts – formed by mapping the pluralities to the abstracta – are identical, because the fusions overlap the same nonstationary – i.e., $\neq \emptyset$ – locations. The converse is the determinative ground of the identification.)

The foregoing constraints on the formality of the parthood relation – both being necessarily non-concrete and figuring in pluralities which serve to individuate fusions as abstract objects – are sufficient then for redressing the objections to the second dogma of semantic rationalism; i.e., that individuation-conditions are wanting for concepts unique to formal and informal domains, which would subsequently render the modal profile of such concepts indeterminate. That relations of mereological parthood are abstract adduces in favor of the claim that the values taken by the relation are necessary. The significance of both the necessity of the parthood relation, as well as its being abstract rather than concrete and thus being in some sense apriori, is that there are thus compelling grounds for taking the relation to be super-rigid, i.e., to be both epistemically and metaphysically necessary.

Finally, a third issue, related to the second dogma is that, following Dummett (1963/1978: 195-196), the concept of mereological parthood might be taken to exhibit a type of 'inherent vagueness', in virtue of being indefinitely extensible. Dummett (1996: 441) defines an indefinitely extensible concept as being such that: 'if we can form a definite conception of a totality all of whose members fall under the concept, we can, by reference to that totality,

---

[5]μ is further Alexandrov; i.e., closed under arbitrary unions and intersections.

characterize a larger totality all of whose members fall under it'. It will thus be always possible to increase the size of the domain of elements over which one quantifies, in virtue of the nature of the concept at issue; e.g., the concept of ordinal number is such that ordinals can continue to be generated, despite the endeavor to quantify over a complete domain, in virtue of iterated applications of the successor relation, and the concept of real number is such that the reals can continue to be generated via elementary embeddings. Bernays' (1942) theorem states that class-valued functions from classes to sub-classes are not onto, where classes are non-sets (cf. Uzquiano, 2015a: 186-187). A generalization of Bernays' theorem can be recorded in plural set theory,[6] where the cardinality of the sub-pluralities of an incipient plurality will always be greater than the size of that incipient plurality. If one takes the cardinal height of the cumulative hierarchy to be fixed, then one way of tracking the variance in the cardinal size falling in the extension of the concept of mereological parthood might be by redefining the intension thereof (Uzquiano, 2015b). Because it would always be possible to reinterpret the concept's intension in order to track the increase in the size of the plural universe, the intension of the concept would subsequently be non-rigid; and the concept would thus no longer be super-rigid.

One way in which the objection might be countered is by construing the variance in the intension of the concept of parthood as tracking temporal modal properties, rather than metaphysical modal properties. Then, the relation can be necessary while satisfying full S5 – i.e., modal axioms K [($\Box \phi \rightarrow \psi) \rightarrow (\Box \phi \rightarrow \Box \psi)$], T ($\Box \phi \rightarrow \phi$), and E ($\neg \Box \phi \rightarrow \Box \neg \Box \phi$) – despite that there can be variations in the size of the quantifier domains over which the relation and its concept are defined. Let $\uparrow$ be an intensional parameter which indexes and stores the relevant formulas at issue to a particular world (cf. Hodes, 1984). The $\downarrow$-symbol is an operator which serves to retrieve, as it were, that indexed information. Adding multiple arrows is then akin to multiple-indexing: The value of a formula, as indexed to a particular world, will then constrain the value of that formula, as indexed – via the addition of the new arrows – to different worlds. Interpreting the operators temporally permits there to be multiple-indexing in the array of intensional parameters relative to which a formula gets its value, while the underlying

---

[6]See Burgess (2004/2008), for an axiomatization of 'Boolos-Bernays' plural set theory, so named after the contributions of Bernays (op. cit.) and Boolos (1984, 1985). See Linnebo (2007), for critical discussion.

logic for metaphysical modal operators can be S5, partitioning the space of worlds into equivalence classes. Formally:

$\uparrow_1 \forall x \exists \phi \uparrow_2 \exists y [\phi(x) \downarrow_1 \wedge \phi(y) \downarrow_2]$.

The clause states that, relative to a first temporal parameter in which all of the x's satisfying the sethood predicate are quantified over, there is – relative to a distinct temporal parameter – another element which satisfies that predicate. Crucially, differences in the intensional temporal indices, as availed of in order to record variance, at different times, in the size of the cumulative hierarchy of elements falling in the range of the parthood relation, is yet consistent with the cardinality of the elements in the domain falling in the range of the relation being fixed, such that the valuation of the relation can yet be necessary.

4.3.4   Summary

In this section, I addressed objections to two dogmas of the semantic rationalism underpinning the epistemic interpretation of two-dimensional intensional semantics. In response to objections to the first dogma – according to which no distinctions can be delineated between linguistic intensions and conceptual epistemic intensions – I noted that higher-order plural terms are conceptually tractable although they have no analogue in natural language semantics. In response to objections to the second dogma – according to which criteria on distinguishing formal from informal domains unique to the extensions of various concepts are lacking, which subsequently engenders indeterminacy with regard to the modal profiles of those concepts – I availed of generic criterial identity conditions, in which it is essential to identical arbitrary representatives of objects that they satisfy equivalence relations which are conversely ground-theoretically determinative of the identification, and further essential thereto that they satisfy the predicate of being necessarily non-concrete. The extensions of indefinitely extensible concepts can further be redefined relative to distinct temporal intensional parameters, despite that the background modal logic for the intensions of the concepts partitions the domain of worlds into equivalence classes, and thus satisfies S5. Thus, parthood can be deemed a necessary, because abstract, relation, despite (i) temporal variance in the particular objects on which the parthood relation is defined; and (ii) variance in the cardinality of the domain in which those objects figure, relative to which the concept's intensions are defined.

When $\Phi = x \preceq xx$, $\Box \forall x, y \Box \exists \Phi (\Phi x \iff x = y)$. By the super-rigidity of

the parthood relation, the target two-dimensional intensional formula can, finally, be stated as follows:

If it is epistemically possible that Φx, then it is metaphysically possible that Φx. Formally:

$\forall c \in C, p \in P[\![\Phi x]\!]^{c,p} = 1$ iff $\forall c' \in C, p' \in P[\![\Phi x]\!]^{c',p'} = 1$.

Thus, the epistemic possibility of haecceity comprehension constrains the value of the metaphysical possibility of haecceity comprehension, and – in response to Roca-Royes and Chalmers – there is a case according to which conceivability is a guide to a principle of modal metaphysics.

In the remainder of the paper, I will examine issues pertaining to the determinacy of epistemic possibilities.

4.4   Determinacy and Consistency

In his (2014), Chalmers argues for the law of excluded middle, such that it is either apriori derivable using the material conditional – i.e. 'scrutable' – that p or scrutable that ¬p, depending on the determinacy of p. Chalmers refers to the case in which p must be determinate, entailing determinate scrutability, as the Hawthorne model, and the case in which it can be indeterminate, entailing indeterminate scrutability, as the Dorr model (259).[7] Chalmers argues that, for any p, one can derive 'p iff it is scrutable that p' from 'p iff it is true that p' (262). However, 'p iff it is scrutable that p' is unrestrictedly valid only on Dorr's, and not Hawthorne's, model (op. cit.).[8]

Chalmers suggests that the relevant notion of consistency might be a property of epistemic possibilities rather than metaphysical possibilities. How-

---

[7]Cf. Dorr (2003: 103-4) and Hawthorne (2005: sec. 2).

[8]Chalmers rejects the epistemicist approach to indeterminacy, which reconciles the determinacy in the value of a proposition with the epistemic indeterminacy concerning whether the proposition is known (op. cit.: 288). Consider, e.g., a color continuum, beginning with a determinate color hue of red and terminating with a determinate color hue of orange. By transitivity, if the determinate hue of red, x, is phenomenally similar to the next point in the continuum, y, and y is phenomenally similar to the next point, z, then x is phenomenally similar to z. However, iterating transitivity would entail that the terminal color hue is red and not orange. Thus, if the culprit in the sorites paradox is the property of transitivity, then the modal axiom which encodes transitivity (namely 4: $\Box\phi \rightarrow \Box\Box\phi$) is false. The epistemic interpretation of the axiom states that if one knows that ϕ, then one knows that one knows that ϕ. Thus, rejecting axiom 4 entails that the cut-off points in a sorites series are knowable, although one cannot know that one knows them. For further discussion, see Williamson (1994).

ever, there are general barriers to establishing the consistency of the space of epistemic modality.

One route to securing the epistemic interpretation of consistency is via Chalmers' conception of idealized epistemic possibility. Conceivability is ideal if and only if nothing rules it out apriori upon unbounded rational reflection (2012: 143). The rational reflection pertinent to idealized conceivability can be countenanced modally, normatively, and so as to concern the notion of epistemic entitlement. An idealization is (i) modal iff it concerns what it is metaphysically possible for an agent to know or believe; (ii) normative iff it concerns what agents ought to believe; and (iii) warrant-involving iff it concerns the propositions which agents are implicitly entitled to believe (2012: 63). It is unclear whether any of (i)-(iii) in the foregoing would either mandate belief in the claim that 'p $\wedge$ it is indeterminate whether p' is true, or explain in virtue of what the conjuncts are consistent. More general issues for the consistency of epistemically possible worlds, even assuming that the idealization conditions specified in (i)-(iii) are satisfied, include Yablo's (1993) paradox, and Gödel's (1931) incompleteness theorems. Yablo's paradox is as follows:

(S1) For all k>1, S_k is false;

(S2) For all k>2, S_k is false;

...

(Sn) For all k>n, S_k is false;

(Sn+1) For all k>n+1, S_k is false.

(Sn) says that (Sn+1) is false. Yet (Sn+1) is true. Contradiction.[9]

Gödel's incompleteness theorems can be thus outlined. Relative to a choice of (i) coding for an ω-complete, recursively axiomatizable language, L – i.e. a mapping between properties of numbers and properties of terms and formulas in L; (ii) a predicate, phi; and (iii) a fixed-point construction: Let phi express the property of 'being provable', and define (iii) s.t., for any consistent theory T of L, there are sentences, p_{phi}, corresponding to each formula, phi(x), in T, s.t. for 'm' := p_{phi},

$\vdash_T p_{phi}$ iff phi(m).

One can then construct a sentence, 'm' := ¬phi(m), such that L is incomplete (the first incompleteness theorem).

Crucially, moreover, L cannot prove its own consistency:

If:

---

[9]For further discussion, see Cook (2014).

$\vdash_T$ 'm' iff ¬phi(m),

Then:

$\vdash_T$ C → m.

So, L is consistent only if L is inconsistent (the second incompleteness theorem).

Another issue concerning the consistency of 'p ∧ it is indeterminate whether p' – let alone the foregoing general issues concerning the consistency of epistemic modal space – is that Chalmers (2009: 102) endorses the indeterminacy of *metaphysical* proposals such as unrestricted fusion and, presumably, the necessity of parthood, with regard to which the epistemic interpretation of consistency would be irrelevant (264).

To redress the issue, the metaphysical indeterminacy of ontological proposals might be treated as in Barnes and Williams (2011), for whom metaphysical indeterminacy consists in there being an unpointed set of metaphysically possible worlds; i.e., a set of metaphysical possibilities, P, such that precisifications concerning the determinacy in the values of the elements of P leave it unsettled which possibility is actual (116, 124). If so, then metaphysical indeterminacy will provide no new objection to the viability of the two-dimensional intensional framework, because the conditions on ascertaining the actuality of the epistemic possibility in the context – relative to which a formula receives a value, and thus crucially determines the value of the formula relative to an index which ranges over metaphysically possible worlds – are themselves indeterminate (cf. Yablo, 2008).

The more compelling maneuver might instead be to restrict the valid apriori material entailments to determinately true propositions; and to argue, against Chalmers's preferred ontological anti-realist methodology, that the necessity of parthood is both epistemically and metaphysically determinately true, if true at all. The (determinate) truth of the proposition might then be corroborated both by the consistency of its augmentation to the logic underlying the intensional semantics, and perhaps in virtue of other abductive criteria – such as strength, simplicitly, and compatability with what is known – on the tenability of the proposal.

4.5   Concluding Remarks

One of the primary objections to accounting for the relationship between conceivability and metaphysical possibility via the epistemic interpretation of

two-dimensional intensional semantics is that epistemic possibilities are purportedly insensitive to modal metaphysical propositions, concerning, e.g., the haecceitistic properties of individuals. In this paper, I have endeavored to redress the foregoing objection, by noting that it relies on an unduly restrictive, propositional view of the elements of epistemic modal space, which ignores the interaction between epistemic possibilities and the nature of higher-order quantification. Further objections, from both the potential indeterminacy in, and inconsistency of, the space of epistemic possibilities, were then shown to be readily answered. In virtue of the super-rigidity of the parthood relation, epistemic modal space can thus serve as a guide to haecceity comprehension principles in modal metaphysics, and thus to *de re* modality.

Chapter 5

Grounding, Conceivability, and the Mind-Body Problem

This paper argues that Chalmers's (1996; 2010) two-dimensional conceivability argument against the derivation of phenomenal truths from physical truths risks being obviated by a hyperintensional regimentation of the ontology of consciousness.

Chalmers (2010) provides the following argument against the identification of phenomenal truths with physical and functional truths. Let M be a model comprised of a domain D of formulas; C a set of epistemic possibilities; W a set of metaphysical possibilities; R_c and R_w, accessibility relations on C and W, respectively; and V a valuation function assigning formulas to subsets of C and W. So, M = $\langle$D,C,W,R_c,R_w,V$\rangle$. Let P denote the subset of formulas in the domain concerning fundamental physics, as well as both neurofunctional properties such as oscillations of neural populations, and psychofunctional properties such as the retrieval of information from memory stores. Let Q denote the subset of formulas in the domain concerning phenomenal consciousness. A formula is epistemically necessary or apriori ($\Box$), if and only if it has the same value at all points in C, if and only if it is impossible, i.e. inconceivable, for the formula to a variant value ($\neg\Diamond\neg$). A formula is negatively conceivable ($\Diamond$) if and only if nothing rules it out apriori ($\neg\Box\neg$) (144). A formula is metaphysically necessary if and only if it has the same value at all points in W. A formula is said to be 'super-rigid', if and only if it is both epistemically and metaphysically necessary, and thus has the same value at all points in epistemic and metaphysical modal space (2012: 474).

The two-dimensional conceivability argument against physicalism proceeds as follows.

The physicalist thesis states that:
P → Q.
Suppose, however, that the physicalist thesis is false. Thus,
1. ¬(P → Q).
By the definition of the material conditional,
2. ¬(¬P ∨ Q).
By the De Morgan rules for negation,
3. ¬¬P ∧ ¬Q.
By double negation elimination,
4. P ∧ ¬Q.[1]

'P ∧ ¬Q' can receive a truth value relative to two parameters, a context, C, and an index, W. In two-dimensional intensional semantics, the value of the formula relative to the context determines the value of the formula relative to the index. Let the context range over a space of epistemic possibilities and let the index range over a space of metaphysical possibilities. Then,

$\llbracket P \wedge \neg Q \rrbracket^{c,w} = 1$ iff $\exists c' \in C \exists w' \in W \llbracket P \wedge \neg Q \rrbracket^{c',w'} = 1$.[2]

---

[1] For the formal equivalence, given the definition of the material conditional, see Chalmers (2010: 169).

[2] For the clause for the two-dimensional intension, see Chalmers and Rabern (2014: 212). Chalmers' informal characterization of the argument proceeds as follows:

1. P ∧ ¬Q is conceivable.
2. If P ∧ ¬Q is conceivable, P ∧ ¬Q is [epistemically, i.e.] 1-possible.
3. If P ∧ ¬Q is 1-possible, P ∧ ¬Q is [metaphysically, i.e.] 2-possible.
4. If P ∧ ¬Q is 2-possible, then materialism is false.
Thus,
5. Materialism is false (2010: 149).
The foregoing clause codifies the thought that, if it is epistemically possible that the truths about physics and functional organization obtain while the truths about consciousness do not, then the dissociation between P and Q is metaphysically possible as well. The argument depends on the assumption that propositions about consciousness and physics are super-rigid, such that the epistemic possibility concerning such truths can serve as a guide to the metaphysical possibility thereof.
If the conceivability argument is sound, then the physicalist thesis – that all phenomenal truths are derivable from physical and functional truths – is possibly false. The foregoing argument entrains, thereby, the metaphysical possibility of a property-based version of dualism between phenomenal consciousness and fundamental physics.
One of the standard responses to Chalmers's conceivability argument is to endeavor to

argue that there are 'strong' necessities, i.e. cases according to which the necessity of the physical and phenomenal formulas throughout epistemic and metaphysical modal space is yet consistent with the epistemic possibility that the formulas have a different value.[3] Note, however, that strong necessities are ruled-out, just if one accepts the normal duality axioms for the modal operators: i.e., it is necessary that ϕ if and only if it is impossible for ϕ to be false: $\Box\phi$ iff $\neg\diamond\neg\phi$. Thus, the epistemic necessity of ϕ rules out the epistemic possibility of not-ϕ by fiat. So, proponents of the strong necessity strategy are committed to a revision of the classical duality axioms.

Another line of counter-argument proceeds by suggesting that the formulas and terms at issue are not super-rigid. Against the super-rigidity of physical truths, one might argue, for example, that our knowledge of fundamental physics is incomplete, such that there might be newly discovered phenomenal or proto-phenomenal truths in physical theories from which the truths about consciousness might be derived.[4] More contentiously, the epistemic profile of consciousness – as recorded by the concepts comprising our thoughts thereof, or by the appearance of its instantiation – might be dissociable from its actual instantiation. A variation on this reply takes our concepts of phenomenal consciousness still to refer to physical properties (cf. Block, 2006). A related line of counter-argument relies on the assumption that phenomenal concepts are entities which are themselves physically reducible (cf. Balog, 1999).

Finally, a counter-argument to the conceivability argument that has yet to be advanced in the literature is that its underlying logic might be non-classical. Thus, for example – by relying on double negation elimination in the inference from line 3 to 4 above – the two-dimensional conceivability argument is intuitionistically invalid. A novel approach might further consist in arguing that epistemic modality might be governed by the Routley-Meyer semantics for relevant logic.[5] Relevant validity can be defined via a ternary relation, such that $[\![\phi \rightarrow \psi]\!]^\alpha = 1$ iff $[\![\phi]\!]^\beta \leq [\![\psi]\!]^\gamma$ and $R(\alpha,\beta,\gamma)$, where the parameters, α, β, and γ, range over epistemic possibilities. Then the irrelevant entailment, $\phi \wedge \neg\phi \rightarrow \psi$, can be avoided by setting $[\![\phi]\!]^\beta = 1$; $[\![\phi]\!]^\gamma = 0$; $[\![\psi]\!]^\beta = 0$; while $[\![\psi]\!]^\gamma = 1$. So, $[\![\phi]\!]^\beta = 1$; $[\![\phi]\!]^\gamma = 0$; and $[\![\psi]\!]^\beta = 0$. The philosophical interest of relevant logic is that it eschews the principle of disjunctive syllogism; i.e., $\forall\phi,\psi[[(\phi \vee \psi) \wedge \neg\phi] \rightarrow \psi]$ and $\forall\phi,\psi[[\phi \wedge (\neg\phi \vee \psi)] \rightarrow \psi]$. Without disjunctive syllogism, logical entailment can no longer be identified with the material conditional, and this would block the derivation of line 2 from line 1 in the two-dimensional conceivability argument.

In this essay, I will pursue a line of argument which is novel and distinct from the foregoing. I argue, in turn, that the conceivability argument can be circumvented, when the relationship between the truths about fundamental physics and the truths about phenomenal consciousness is analyzed in a classical, hyperintensional setting. Suppose, for example, that the physicalist thesis is defined using hyperintensional, grounding operators rather than metaphysical necessitation.[6] Then, the epistemic and metaphysical possibility that $\neg(P \rightarrow Q)$ is classically valid, although targets a less fine-grained metaphysical connection between physical and phenomenal truths. Even if P's grounding Q still entails the metaphysical necessitation of Q by P, the epistemic-intensional value of '$\neg(P \rightarrow Q)$' – will be an insufficient guide to the metaphysical-hyperintensional value of the proposition. So, even if the intension for 'consciousness' is rigid in both epistemic and metaphysical modal space, the epistemic intension recording the value of the proposition will be blind

to its actual metaphysical value, because the latter will be hyperintensional.

In the remainder of this essay, I will outline the regimentation of the proposals in the ontology of consciousness using hyperintensional grounding operators, rather than the resources of modality and identity.[7] By contrast to the modal approach underlying the conceivability argument, the hyperintensional regimentation targets the properties of reflexivity and bijective mappings, in order to countenance novel, ontological dependence relations between the properties of consciousness and physics, which are finer-grained than necessitation.[8]

Following Fine (2012a; 2012b), let a polyadic operator have a *ground-theoretic* interpretation, only if the profile induced by the interpretation concerns the hyperintensional *truth-making* connection between an antecedent set of truths or properties and the relevant consequent. Let a grounding operator be *weak* if and only if it induces reflexive grounding; i.e., if and only if it is sufficient for the provision of its own ground. A grounding operator is *strict* if and only if it is not weak. A grounding operator is *full* if and only if it uniquely provides the explanatory ground for a fact. A grounding operator is *part* if and only if it - along with other facts - provide the explanatory ground for a fusion of facts.

Combinations of the foregoing explanatory operators may also obtain: $x < y$ iff ϕ is a *strict full* ground for ψ; $x \leq y$ iff ϕ is a *weak full* ground for ψ; $x \prec y$ iff ϕ is a *strict part* ground for ψ; $x \preceq y$ iff ϕ is a *weak part* ground for ψ; $x \preceq y \wedge \neg(y \preceq x)$ iff ϕ is a *strict partial* ground for ψ; $x \prec^* y$ iff $x_1, ..., x_n \leq y$, iff ϕ is a *partial strict* ground for ψ; $x \prec' z$ iff $[\phi \prec^* \psi \wedge \psi \preceq \mu]$ iff ϕ is a *part strict* ground for some further fact, μ.[9]

The proposals in the metaphysics of consciousness can then be regimented in the hyperintensional framework as follows.

- Functionalism (modally: truths about consciousness are identical to truths about neuro- or psychofunctional role):

  Functional truths (F) ground truths about consciousness (Q) if and only if the grounding operator is:

  -strict full, s.t. $F < Q$

  -distributive (i.e. bijective between each truth-ground and grounded truth), s.t. $\exists f_{1-1}\langle F, Q\rangle$

- Phenomenal Realist Type Identity (modally: truths about consciousness are identical to truths about biological properties, yet phenomenal properties are – in some sense – non-reductively real).[10]

  Biological truths (B) ground truths about consciousness (Q) if and only if the grounding operator is:

  -strict partial, s.t. $B \preceq Q \wedge \neg \, Q \preceq B$;

  -distributive, s.t. $\exists f_{1-1}\langle B, Q\rangle$; and

  -truths about consciousness are weak part (i.e. the set partly reflexively grounds itself), s.t. $Q \preceq Q$

- Property Dualism (modally: truths about consciousness are identical neither to functional nor biological truths, yet are necessitated by physical truths):

Physical truths (P) ground truths about consciousness (Q) if and only if the grounding operator is:

-P $\preceq$ Q;

-non-distributive, s.t. $\neg \exists f_{1-1} \langle$P, Q$\rangle$; and

-truths about consciousness are weak part, s.t. Q $\preceq$ Q

- Panpsychism (in Non-constitutive guise: Phenomenal properties are the intrinsic realizers of extrinsic functional properties and their roles; in Constitutive guise: (i) fundamental microphysical entities are functionally specified and they instantiate microphenomenal properties, where microphenomenal properties are the realizers of the fundamental microphysical entity's role/functional specification; and (ii) microphenomenal properties constitute the macrophenomenal properties of macrophysical entities):

    Truths about consciousness (Q) ground truths about functional role (F) if and only if the grounding operator is:

    -strict full, s.t. Q < F; and

    -non-distributive, s.t. $\neg \exists f_{1-1} \langle$Q, F$\rangle$

The philosophical significance of the hyperintensional regimentation of the ontology of consciousness is at least three-fold. First, the regimentation permits one coherently to formulate Phenomenal Reality Type Identity. Leibniz's law states that for all propositional variables x,y and for all properties R, x = y iff (Rx $\iff$ Ry). According to the Phenomenal Realist Type Identity proposal, phenomenal properties are identicial to biological properties, while phenomenal properties are in some sense non-reductively real. Thus, in the modal setting, Phenomenal Realist Type Identity belies Leibniz's law, on the assumption that the latter can be applied to intensional entities. One virtue of the hyperintensional regimentation is thus that it avoids this result, by providing a framework with the expressive resources sufficient to formulate the non-reductive Type Identity proposal.

Second, the hyperintensional grounding regimentation evinces how functionalist approaches to the ontology of consciousness can be explanatory, because the identification of phenomenal properties with functional organization can be defined via the foregoing ground-theoretic explanatory properties. Block (2015) suggests that – by contrast to Phenomenal Realist Type Identity – identifying phenomenal properties with functional roles cannot sufficiently account for the ground-theoretic explanation of the identity. Block distinguishes between metaphysical and ontological versions of physicalism. Block's 'ontological physicalism' is a reductive, functionalist theory, and eschews of explanation by restricting the remit of its theory to 'what there is'; i.e. to specifying identity statements between entities in the domain of quantification (114). By contrast, Block's 'metaphysical physicalism' – namely, Phenomenal Realist Type Identity – purports to account for the nature of the entities figuring in theoretical identity statements via availing of relations of explanatory, ontological dependence (op. cit.).

Block poses the following consideration against the functionalist (117). Suppose that there is a counterpart of a human organism with isomorphic functional properties, but comprised of distinct biological properties. Suppose that the functional isomorph instantiates phenomenal properties. Block argues that the functional isomorph 'is like us

superficially, but not in any deep property that can plausibly be one that scientists will one day tell us is the physical ground of consciousness [. . .] So there is a key question that that kind of reductive physicalism – ontological physicalism – does not ask nor answer: what is it that creatures with the same phenomenology share that grounds that phenomenology' (op. cit.)? The foregoing does not provide an argument that the neuro- and psycho-functionalist must provide an account of in virtue of what phenomenal properties are instantiated. Rather, Block suggests only that functionalist proposals do not sufficiently inquire into the realizers of the functional roles that they specify. He suggests that this theoretical approach would be insufficient, if one were to seek an explanation of the psychofunctional correlations between phenomenal property types and the relevant functional roles.

The second theoretical virtue of the hyperintensional regimentation is thus that it demonstrates how Block's analysis might be circumvented. Functionalism can be regimented within the logic of hyperintensional ground; and can therefore satisfy the formal requirements on explaining in virtue of what phenomenal truths ontologically depend upon functional truths [cf. Khudairi (op. cit.)].[11]

Third, and most crucially: The regimentation demonstrates how metaphysically possible relations between consciousness and physics cannot be witnessed by epistemic constraints, when the latter are recorded by the conceivability – i.e., the epistemic possibility – thereof. Propositional epistemic modality is blind to the hyperintensional, metaphysical dependencies holding between phenomenal and physical truths. Thus, the two-dimensional conceivability argument against the derivation of phenomenal truths from physical truths risks being obviated by a hyperintensional regimentation of the ontology of consciousness.

Chapter 6

Attention, Phenomenal Content, and Inexact Knowledge

6.1 Introduction

This chapter aims to precisify the relations between the phenomenological effects of attention and the veridicality-conditions of perceptual states. Bayesian perceptual psychology provides a model which explains how perceptual states can be understood as possessing conditions of accuracy or veridicality. Once unconscious perceptual representational states have been derived by the visual system, a further issue concerns the conditions that are necessary and sufficient for the perceptual states to become phenomenally conscious. The most promising candidate for at least a necessary condition on the instantiation of phenomenal properties on perceptual representational states is the operation of the mechanisms of attention. Problematically, however, the distribution of attentional mechanisms has been confirmed as entraining an additional effect, according to which attention to features or properties – such as color-saturation, contrastivity, gap size, and the size of moving stimuli – significantly changes the perceived value of the target property. The effect is prevalent, computationally tractable, and systematic. Thus, one crucial question concerns whether the attentional effects entail that the perceptual states are illusory. If so, then the effects of attention in changing the perceived value of phenomenal properties raises a dilemma

concerning whether the correctness-conditions on perceptual states are ever normally satisfied.

In this chapter, I argue that the foregoing dilemma is not robust. Four issues must be distinguished: (i) the first issue concerns the accuracy-conditions of unconscious perceptual states; (ii) the second issue concerns the accuracy-conditions of phenomenal properties; (iii) the third issue concerns how attention alters the perceived value of phenomenal properties, once the perceptual states have become conscious; and (iv) the fourth issue concerns how the accuracy-conditions of perceptual states interact with the concept of knowledge. Issue (i) is well-understood.[1] There are promising approaches to issue (ii).[2] Issue (iii) has further been carefully documented.[3] The remaining issue is therefore issue (iv); i.e. the question of how the satisfaction of the constitutive-conditions for perceptual representations interacts with the concept of knowledge. The concept of knowledge – and of phenomenal knowledge, in particular – has been examined and made precise by availing of epistemic modal logic, with applications to subjective probability and vagueness. I argue, then, that the concept of knowledge is not immediately relevant to issue (i): Although they interact in philosophically significant ways, Bayesian perceptual state space is distinct from epistemic logic. The remaining question concerns, then, the epistemic status of the properties of phenomenal consciousness, in virtue of the value-altering functions of attention. I argue that phenomenal properties are low-level properties – such as color-hue, shape, and orientation – and that accuracy-conditions for the properties can be constructed via a category-theoretic semantics. When attention alters the perceived value of a phenomenal property, then the typed arrow that represents the property might subsequently not be able to figure in the composition of the concatenation of arrows which ultimately maps to the values, [0,1], in the co-domain. However, that result is consistent with the value of the phenomenal property still being known, and thereby factive. Thus, the phenomenological effects of attention affect neither the knowledge nor the factivity of the phenomenal properties at issue; and a precise explana-

---

[1]For the loci classici, see Egan (2010) and Burge (2010). See Section **2**, for further discussion.

[2]See, inter alia, Dretske (1995); Tye (2000); Byrne (2009); Siegel (2010); and Chalmers (2010). Chalmers' approach is targeted, below, in Section **3**.

[3]For a survey of the empirical results relevant to the phenomenological effects of attention, see Carrassco (2011). For the locus classicus with regard to the philosophical significance of the results, see Block (2010). See Section **4**, for further discussion.

tion can be provided of the manner in which the attentional effects indirectly bear on the contents of phenomenal consciousness.

The chapter has, accordingly, two aims. The first is to provide a precise account of the manner in which the properties of phenomenal consciousness can be understood to possess accuracy-conditions. The present account will be limited to examining sensory phenomenal properties. Whether phenomenal properties can also be unique to each of cognitive, conative, emotional and ethical attitudes is a topic requiring separate treatment. The second aim of the chapter is to examine how content relates to knowledge.

In Section **2**, I detail the constitutive conditions for perceptual representational states. In Section **3**, I provide a critical discussion of an approach to phenomenal accuracy-conditions, which avails of the epistemic interpretation of two-dimensional intensional semantics. I examine several general issues for the the foregoing approach, which I argue adduce against the application of the semantic framework to phenomenal representation. In turn, I provide a novel proposal of the veridicality-conditions for phenomenal properties. I demonstrate how the proposal is empirically adequate, by constitutively incorporating both the empirical results concerning the kinds of attention and the model-theoretic properties availed of in vision science. I aim, thereby, to provide a formally tractable account of the manner in which phenomenal properties can precisely compose via a category-theoretic semantics; and I argue that this concatenation of phenomenal properties into propositional form – assessable for accuracy or veridicality – explains the nature of the correctness-conditions for phenomenal consciousness. In Section **4**, I examine how the formal properties of content relate to the logic of knowledge. I argue that the phenomenological effects of attention affect neither the satisfaction of phenomenal knowledge nor the factivity thereof, and that the value-altering effects of attention bear on phenomenal content only by targeting its component parts, in a manner similar to how changes in the semantic values of sub-sentential expressions might affect the truth-condition for the sentence that the component expressions comprise. Section **5** provides concluding remarks.

6.2   Perceptual Representation

In order to explain the visual system's derivation of representation-conditions, the model availed of in contemporary vision science is a Bayesian frame or

state space. (I will prescind, in what follows, from specifying a valuation function for spaces, and focus just on frames.) The frame is Bayesian, because the probability measures defined on the frame satisfy the Kolmogorov axioms.[4] In a Bayesian frame, $\langle \Omega, Pr \rangle$, the states at issue are possibilities concerning the source of retinal, sensory information. The frame is closed under complementation and intersection, such that a σ-algebra is defined thereon. Crucially, the frame is pointed. The visual system computes which of the possibilities in the state frame has the highest probability of being actual.[5] In the frame of possibilities, the state with the highest probability of being actual is referred to as the constancy. The constancies – of, e.g., lightness and distance – comprise accuracy-conditions on the attribution of spatial properties – e.g., volume and boundedness – to distal particulars. The constancies are thus identified with the constitutive conditions of unconscious perceptual representations.

It is a point of contention whether the representational properties entrained by the visual system's computations entail that the computational properties are constitutively representational (cf. Burge, op. cit.; Rescorla, 2013), or are only subsequently interpreted as specifying content (cf. Egan, op. cit.; Chalmers, 2012). However, it is sufficient for the purposes of this note that representational properties are at least entailed by the visual system's computational properties – i.e., the derivation of the constancies – as modeled above.

In the following section, I will discuss two approaches to specifying the accuracy-conditions for the properties of phenomenal consciousness, by contrast to those for unconscious perceptual states. The first avails of two-dimensional semantics on its epistemic interpretation. I target several, general issues for the two-dimensional approach, and – in turn – propose a novel account of the accuracy-conditions which is based upon a category-theoretic semantics.

---

[4]The axioms are normality – 'Pr(T) = 1'; non-negativity – 'Pr(ϕ) > 0'; additivity – for disjoint ϕ and ψ, 'Pr($\phi \cup \psi$) = Pr(ϕ) + Pr(ψ)'; and conditionalization – 'Pr($\phi|\psi$) = Pr($\phi \cap \psi$) / Pr(ψ)'.

[5]Cf. Mamassian et al (2002).

6.3 Phenomenal Representation

6.3.1 The Two-stage Theory of Phenomenal Content

According to the two-dimensional intensional approach alluded to above, a distinction can be drawn between perfect and imperfect correctness-conditions for phenomenal properties. The approach has accordingly been termed the 'two-stage' theory (cf. Chalmers, 2010: 402-418). I will argue that there are reasons to believe that the theory is conceptually problematic and empirically implausible. I will then proceed to outline a positive proposal of the accuracy-conditions for phenomenal properties, and I argue that the proposal is able to circumvent the issues adumbrated for the two-stage account.

According to the two-stage theory, a relation of matching is stipulated to exist between perfect and imperfect properties. The accuracy-conditions for perfect properties are epistemically possible, although not metaphysically actual (413). Via conceivability or exercises of the imagination, individuals are acquainted with perfect accuracy-conditions for possible phenomenal properties (op. cit.). The correctness-conditions are thus individuated by intensions defined on a space of epistemically possible worlds. Perfect contents are referred to as 'Edenic' (op. cit.: 403), although – because of the role of conceivability in acquaintance with such accuracy-conditions – I will occasionally contrive to deploy the more evocative, 'Transcendent' (`Tr`). Transcendent acquaintance with non-actual, possibly satisfiable accuracy-conditions for perfect phenomenal properties is thus purported to admit of phenomenal *presentation* (406). By contrast, actually satisfiable, ordinary accuracy-conditions for the imperfect, instantiated phenomenal properties – sullied by noise and illusion – are claimed to constitute phenomenal *representation* (op. cit.).

Ordinary phenomenal representations are themselves claimed to take two distinct forms, Fregean (`Fr`) and Russellian (`Rs`). `Fr` is satisfiable if and only if 'its object has the property that normally causes a phenomenally red experience' (403). `Rs` is satisfiable if and only if the accuracy-condition is 'object-involving', s.t. 'it involves a specific object O, and its satisfaction depends on the properties of O' (385), if and only if 'its object has physical redness' (403).

Why stipulate that there are at least two types of accuracy-condition: the ordinary and imperfect, by contrast to the transcendent and perfect? Chalmers argues that 'phenomenally red experience is imperfectly veridical

iff its object instantiates a property that matches perfect redness' (404-5). He avers that perfect accuracy-conditions – although epistemically and meta-phyiscally possible but non-actual – are 'fundamental' (405, 413). Thus, as a transcendent though non-actually possible 'regulative ideal', a perfect accuracy-condition 'determines the ordinary [`Fr` accuracy-condition] of the experience' (405-406), where the `Fr`-condition itself bears a relation to the `Rs`-condition.

The general framework for the foregoing proposal is the epistemic inter-pretation of two-dimensional intensional semantics. According to the ap-proach, the counterfactual assessment of the values taken by a term or for-mula is determined by the value antecedently assigned to the term or for-mula in an epistemically possible world, ascertained by various means as being actual. Similarly, with regard to the foregoing 'two-stage', pluralis-tic approach to the accuracy-conditions of phenomenal consciousness: (i) a function, '$\forall x. f_{Tr}(x)$', is specified in an ideal or vindicated, non-actual epis-temically possible world, mapping terms and formulas therein to; (ii) a dis-tinct set of functions, $f_{Fr}(x)$ – where the latter are intensions, i.e., functions from a set of epistemic worlds designated as actual to extensions; and (iii) the extension of the intension, $f_{Fr}(x)$, is then f_{Rs}, i.e. itself a function from metaphysically possible worlds to extensions.

The foregoing proposal is coherent, and possesses the virtue of simulta-neously accommodating several kinds of accuracy-condition for phenomenal properties. However, there are crucial issues for the approach that merit mention.

The first issue concerns the empirical adequacy of the transcendental method associated with the epistemic interpretation of two-dimensional in-tensional semantics. By contrast to this aprioristic, top-down method, an empirically adequate theory of accuracy-conditions might instead (i) demand a (non-reductive) naturalistic explanation, in which the relevant conditions of empirical adequacy first target the accuracy-conditions of unconscious per-ceptual representations, as augmented by specifying the associated biological property types and functional properties such as those of attention; and (ii) only subsequently examine the instantiation- and correctness-conditions of the phenomenal properties that are instantiated thereon. There is thus a sharp distinction between:

(*) the transcendent method of stipulating *ab initio* that there are con-ceivable accuracy-conditions for non-actual perfect properties – taking the latter to be explanatorily fundamental and apriori determining of ordinary

accuracy-conditions for phenomenal states; and

(**) the abductive method of endeavoring first to specify the interaction between biological properties, functional organization, and unconscious representation, and then examining how such empirical properties might constitute phenomenal representation.

Consistently with naturalistic conditions on theory choice, I will pursue, in Section **4** below, the method detailed on alternative (**).[6]

Chalmers argues that a perfect, conceivable accuracy-condition

(i) 'determines the ordinary content of the experience' (405-406), and that

(ii) 'phenomenally red experience is imperfectly veridical iff its object instantiates a property that matches perfect redness' (404-5). To assess the consistency of these claims, we must be more precise.

What are the relata of the matching-relation? The quotation cited in (ii) claims that an ordinary phenomenal property possesses an imperfect correctness-condition if and only if the ordinary phenomenal property matches a non-actual conceivable perfect phenomenal property. The quotation in (i) claims that ideally conceivable accuracy-conditions for perfect properties in some way determine the ordinary accuracy-conditions of imperfect properties. So, by (ii), matching relates phenomenal properties, where matching between perfect and imperfect properties *grounds* imperfect content. By (i), perfect content *determines* imperfect content. So, imperfect content both is determined by perfect content, and holds in virtue of the matching-relation between perfect and imperfect properties.

How can a set of accuracy-conditions determine a distinct set of accuracy-conditions? If 'determine' in this setting is regimented as concerning functions on intensions, then the proposal might be innocuous, modulo the empirical adequacy of the suggestion.

How can the matching-relation between a non-actual epistemically possible phenomenal property and a metaphysically actual phenomenal property explain in virtue of what an accuracy-condition is specified? What would appear to be required is an account of how the matching-relation between distinct properties (perfect and imperfect) explains the correctness-conditions

---

[6]The present, abductive approach may be compared to the 'method of phenomenal contrast' advanced by Siegel (op. cit.). By contrast to the method of phenomenal contrast, which examines the contrasting etiology between two phenomenal properties, the present approach first targets the constitutive conditions of unconscious perceptual representational states, and then the conditions necessary and sufficient for the instantiation of phenomenal properties on such states.

of one of the properties (imperfect) (op. cit.: 404). If imperfect content is determined by perfect content, then it is unclear what could be gained by endeavoring to explain in virtue of what the imperfect content is further specified, and availing of some relation between two properties of which the target content is relevant to only one of the properties.

What Chalmers does note about the proposed matching-relation is that it is a function interpreted as expressing a causal dependence relation (430).[7] However, even if one were to concede for the sake of argument that causation ought generally to back explanation, it remains unclear how a causal relation between distinct properties can explain the correctness-conditions for one of them. Yet this is precisely what is claimed in the passage: 'phenomenally red experience is imperfectly veridical iff its object instantiates a property that [causally] matches perfect redness' (404-5).

A final issue for the proposal concerns the very idea of the causal interpretation of the matching-relation, as applied to perfect and imperfect phenomenal properties. Perfect phenomenal properties are ideally conceivable but not-actual, and thus not actually instantiated. Imperfect phenomenal properties are actually instantiated. How then can a non-instantiated property bear a causal relation to an instantiated property?[8]

In the following section, I will proffer a positive proposal, which aims to circumvent the issues examined above. The proposal that I will develop: (i) targets and explains the accuracy-conditions for phenomenal properties; and (ii) is empirically adequate, by being consistent both with the functional and phenomenological effects of attention and with the Bayesian model theory availed of in vision science.

6.3.2 Category-theoretic Semantics for Phenomenal Representation

In this section, I provide a precise account of the manner in which phenomenal properties can be formally concatenated via a category-theoretic semantics, in order to take propositional form. The general idea behind the semantics is that phenomenal properties are represented by typed arrows, or functions. Typed arrows can thereby be concatenated, via complex predication, and the

---

[7]'[A] property matches a given perfect property (for a given subject) if the property is the normal cause of the associated phenomenal property (in that subject)' (op. cit.).

[8]The objection is owing to Joshua Habgood-Coote (p.c.).

concatenation of the arrows map to the values, $\{0,1\}$, in the co-domain. The precise conditions for the category-theoretic semantics are as follows. With $\vdash$ encoding a representation relation,

$c \vdash$ 'being in pain' iff '$i_c \to \alpha_i$', and

$c \vdash$ 'x is in pain' iff $[\beta] \bullet \alpha_i$: '$i_c \to b$'

iff

$c \vdash$ 'being in pain' denotes f_1, and

$c \vdash$ 'x is in pain' denotes $[[\beta] \bullet f_1](x)$, where

b: $(z^n \to z) = [\beta] = {}'(i_c \in C)^n \to i_c{}'$ as a T-function

$= \langle \, \rangle \to \texttt{True}$, iff $z = \{0,1\}$, s.t.

$b_0(\texttt{F})$ if $z = 0$, and

$b_1(\texttt{T})$ if $z = 1$.

$\blacksquare_{Int} = \text{Int}_C \; [\![\beta f_1]\!] \subseteq [C \subseteq \Omega]$, s.t.

$\beta f_1 \to \texttt{True}$

iff

$M \Vdash \beta f_1$ iff $[\![\beta f_1]\!] = i_c \in C$, and

$M \Vdash i_c \iff i_c \subseteq C$.

Attentional processing has been demonstrated as being a necessary condition on the instantiation of phenomenal properties. Of the kinds of attention, however, the deployment of object-based attention has been challenged as a necessary condition, because of cases in which the minimal resolution of conscious object-based perception is greater than the minimal resolution of object-based attention; such that more objects can be consciously perceived than attended to (cf. Block, 2013). The necessity of endogenous, or top-down, attention has been challenged, as well, conditional on the localization of the fundamental neural correlates of consciousness to occipito-thalamic pathways (cf. Tsuchiya et al, 2012). However, exogenous attention, property-based attention, and spatial-based distributions of attention remain plausible candidates for a necessary condition on the instantiation of phenomenal properties. The attentional mechanisms at issue can be defined psychofunctionally, by their role both in unifying the perception of distinct properties, and in making information available to working memory. The attentional mechanisms at issue can further be defined neurofunctionally, as measured by timed, firing-rate oscillations of neural populations.

Taking stock: I have sought to provide a critical examination of the two-stage theory of phenomenal representation, and subsequently endeavored to outline a positive, alternative proposal. Phenomenal representations can be specified via the category-theoretic construction outlined above, and the

approach was shown both to retain the virtues, and to redress the pitfalls, unique to the two-stage approach.

6.4   Attention and Inexact Knowledge

In this section, I examine, finally, how attention bears on the accuracy-conditions of phenomenal properties, and how phenomenal content is related to phenomenal knowledge.

Phenomenal content interacts with epistemic modal space in at least two ways, depending on whether epistemic modality is interpreted as concerning states of knowledge (the box operator) or belief (the diamond operator), or whether epistemic modality is interpreted as concerning apriority (box) or conceivable scenarios (diamond). In what follows, I will focus on the logic of knowledge and belief, although the interpretation which targets conceivability and apriority merits further discussion.

The foregoing category-theoretic semantics is governed by the modal system, S4. S4 is obtained by the following axioms: K - '$\Box(\phi \to \psi) \to (\Box\phi \to \Box\psi)$'; T - '$\Box\phi \to \phi$'; and 4 - '$\Box\phi \to \Box\Box\phi$'. The topological interpretation is sound and complete for propositional, first-, and higher-order S4 (cf. McKinsey and Tarski, 1944; Awodey and Kishida, 2007; Lando, 2010; and Awodey et al., 2014). The logic of knowledge has similarly been shown at least not to satisfy S5, on pain of sanctioning decidability for the consistency of the subject's epistemic theory of the world. It would be a virtue of the subject's epistemic theory, if the theory's consistency were decidable. However, if the theory were decidable, then it would entail its own consistency, in tension with Gödel's second incompleteness theorem to the effect that, for a first-order, consistent, recursively axiomatizable language of arithmetic in which a sentence attributing the property of unprovability to itself can be defined, the language is consistent only if it is inconsistent (cf. Williamson, 2009).[9] Thus, to avoid the condition that the subject's consistent, epistemic theorizing exceeds the computational power of a Turing machine, the modal system governing the subject's epistemic theory cannot validate the conjunction of axioms 4 and E, i.e. S5 and the sublogics thereof (Williamson, op. cit.). Counterexamples to the transitivity of knowledge have also been witnessed

---

[9] For further discussion of the provability of consistency, and the justification to believe that the background language of the theory is consistent, see Wang (1974: 317-321); Reinhardt (1986: 465-472); Wright (1995); and Detlefsen (1995: 116).

by the sorites paradox and by the phenomena of improbable knowing.[10] In the sorites paradox, suppose that there is a bounded continuum, beginning with a determinate hue of orange and terminating with a determinate hue of red. Transitivity in the model would entail that if one of the bounds of the interval is a hue of orange, then – by transitivity – a matching relation defined on the subsequent points would entail that the remainder of the continuum is orange, contrary to the hypothesis. Thus, transitivity is argued not to hold in the model. On the epistemic interpretation of indeterminacy, if transitivity does not hold, then knowledge of one's epistemic status does not hold, although one's state of knowledge itself – as represented by a sole instance of the box operator – remains unscathed. The example provides, then, an instance of epistemic indeterminacy. Generalizing, one can construct a model according to which transitivity does not hold. Let M_k be a model, s.t. $W_k = \{a,b,c\}$; $a_k = \{a,b\}$; $b_k = \{a,b,c\}$; $c_k = \{b,c\}$; $V_k(\phi)$, ϕ a closed formula $= \{a,b\}$. Thus, $\langle M_k, a \rangle \Vdash \Box\phi$; but not $\langle M_k, b \rangle \Vdash \Box\phi$. So, it is not the case that $\langle M_k, a \rangle \Vdash \Box\Box\phi$; so transitivity does not hold in the model.

Thus, while the topological interpretation of the logic of phenomenal content is S4, the logic of knowledge is KT. Given the discrepancy between the modal systems governing, respectively, the logic of knowledge and the calculus of categories for phenomenal content, the foregoing provides a preliminary argument that phenomenal knowledge is dissociable from phenomenal content; and that the former is more fundamental – because more computationally basic –than the latter.

How, then, do the phenomenological effects of attention bear on the accuracy-conditions for phenomenal properties? Phenomenal properties fall into two types. The first type of phenomenal property is the property of awareness of one's perceptual states. The second type of phenomenal property targets the secondary qualities; for example, pain, color-saturation, and other low-level properties. Because the accuracy-conditions are computationally derived, it would thus be mistaken to argue that attention alters the experience of unconscious perceptual states. Rather, attention alters the perceived value of low-level phenomenal properties, such as color-saturation and gap size.

Block (2010; 2015) has argued for a conclusion similar to the foregoing: no misrepresentation is entrained when attention alters the perceived value of a property. However, in the present framework, the proposal can be extended.

---

[10]See Williamson (1994) and Williamson (2014), for further discussion.

Block's proposal converges with the idea that phenomenal properties – even if not being assessable for accuracy – can still be known. So, the instantiation of phenomenal properties, as modulated by the phenomenological effects of attention, is not illusory, and is instead factive. The foregoing suggests that, in the phenomenal setting, a constraint similar to Frege's context principle is belied. Thus, to ask how the value-altering effects of attention bear on phenomenal properties is comparable to inquiring into the semantic values of various lexical items. How variation in the semantic values bears on the entire truth-condition for the sentence of which the component expressions are only a part – and, similarly, how variation in the values of phenomenal properties bears on the concatenation of the proxies for the properties via the category-theoretic semantics outlined above – is a separate, and not entirely contentious, issue. That the expressions have semantic values, and that phenomenal properties can also have various values, can be known and is thus factive. If so, then the variable effects of attention target only the component parts which comprise the more complex states that are assessable for truth or accuracy.

6.5 Concluding Remarks

In this chapter, I have endeavored to provide a precise account of the accuracy-conditions of phenomenal consciousness, and to explain how the phenonomeno-logical effects of attention interact with both phenomenal content and phenomenal knowledge. Instantiation of a phenomenal property is consistent with knowledge thereof, even if the perceived value of the property is altered in virtue of the phenomenological effects of attention. I argued, finally, that phenomenal properties do not themselves have accuracy-conditions, but are only assessable for veridicality via a form of complex-predication that can be captured by the category-theoretic framework. The derivative effects of attention on the accuracy-conditions of phenomenal consciousness are thus consistent with the inexact knowledge, and thus the factivity, thereof.

Chapter 7

Non-Transitive Self-Knowledge: Luminosity via Modal μ-Automata

This essay provides a novel account of self-knowledge, which avoids the epistemic indeterminacy witnessed by the invalidation of modal axiom 4 in epistemic logic; i.e. the KK principle: $\Box\phi \rightarrow \Box\Box\phi$. The essay argues, by contrast, that – despite the invalidation of modal axiom 4 on its epistemic interpretation – states of epistemic determinacy might yet be secured by countenancing self-knowledge on the model of fixed points in monadic second-order modal logic, i.e. the modal μ-calculus.

Counterinstances to modal axiom 4 – which records the property of transitivity in labeled transition systems, i.e., the relational semantics for modal logic[1] – have been argued to occur within various interpretations of the sorites paradox. Suppose, e.g., that a subject is presented with a bounded continuum, the incipient point of which bears a red color hue and the terminal point of which bears an orange color hue. Suppose, then, that the cut-off points between the points ranging from red to orange are indiscriminable, such that the initial point, a, is determinately red, and matches the next apparent point, b; b matches the next apparent point, c; and thus – by transitivity – a matches c. Similarly, if b matches c, and c matches d, then b matches d. The sorites paradox consists in that iterations of transitivity would entail that the initial and terminal points in the bounded continuum

---

[1]Cf. Kripke (1963).

are phenomenally indistinguishable. However, if one takes transitivity to be the culprit in the sorites, then eschewing the principle would entail a rejection of the corresponding modal axiom (4), which records the iterative nature of the relation. Given the epistemic interpretation of the axiom – namely, that knowledge that a point has a color hue entails knowing that one knows that the point has that color hue – a resolution of the paradox which proceeds by invalidating axiom 4 subsequently entrains the result that one can know that one of the points has a color hue, and yet not know that they know that the point has that color hue (Williamson, 1990: 107-108; 1994: 223-244; 2001: chs. 4-5). The non-transitivity of phenomenal indistinguishability can then provide a structural barrier to higher-order knowledge of one's first-order states. The foregoing result holds, furthermore, in the probabilistic setting, such that the evidential probability that a proposition has a particular value may be certain – i.e., be equal to 1 – while the iteration of the evidential probability operator – recording the evidence with regard to that evidence – is yet equal to 0. Thus, one may be certain on the basis of one's evidence that a proposition has a particular value, while the higher-order evidence with regard to one's evidence adduces entirely against that valuation (Williamson, 2014).

The argument eschews 'safety' as a necessary condition on knowledge, for which Williamson's (2001) approach explicitly argues and as codified by margin-for-error principles of the form: $\forall x \forall \phi [K^{m+1}\phi(x) \rightarrow K^m\phi(x+1)]$' (Williamson, 2001: 128; Gómez-Torrente, 2002: 114); i.e., that if one knows – relative to a margin which ranges over a world accessible from the actual world, m – that an object satisfies a property, then a distinct similar object satisfies that property in the actual world. Intuitively, the safety condition ensures that if one knows that a predicate is satisfied, then one knows that the predicate is satisfied in relevantly similar worlds. Williamson targets the inconsistency of margin-for-error principles, the luminosity principle ['$\forall x \forall \phi [\phi(x) \rightarrow K\phi(x)]$'], and the characterization of the sorites as occurring when an object satisfies a property, such that similar objects would further do so. The triad evinces, arguably, that the safety condition is not satisfied in the sorites, s.t. knowledge does not obtain, and the luminosity principle is false. In cases, further, in which conditions on knowledge are satisfied, epistemic indeterminacy is supposed to issue from the non-transitivity of the accessibility relation on worlds (1994: 242).

One of the primary virtues of the present proposal is thus that it targets the property of transitivity directly, because transitivity both engenders the

sorites paradox on the assumption that the states are known and the property is codified by the epistemic modal axiom for transitivity, i.e., 4 or the KK principle. By so doing, it permits a uniform interpretation of transitivity in the sorites – as codified by the KK principle – such that it applies not only to epistemic accessibility relations whose obtaining is relevant to the safety condition, but further to the logical property and its explanatory role in engendering the paradox.

A second virtue adducing in favor of the foregoing, 'epistemicist' approach to vagueness – which takes the latter to be a phenomenon of epistemic indeterminacy – is that vagueness can be explained without having to revise the underlying logic. The epistemicist approach is consistent with classical logical laws, such as e.g. the law of excluded middle; and thus it can determinately be the case that a point has a color hue; determinately be the case that the next subsequent point has a distinguishable color hue; and one can in principle know where in the continuum the cut-off between the two points lies – yet vagueness will consist in the logical limits – i.e. the non-transitivity – of one's state of knowledge. Thus, one will not in principle be able to know that they know the point at which the color hues are dissimilar.

In this essay, I endeavor to provide a novel account which permits the retention of both classical logic as well as a modal approach to the phenomenon of vagueness, while salvaging the ability of subjects to satisfy necessary conditions on self-knowledge. I will argue that – despite the invalidity of modal axiom 4, given the non-transitivity of the similarity relation – a distinct means of securing an iterated state of knowledge concerning one's first-order knowledge that a particular state obtains is by availing of fixed point, nondeterministic automata in the setting of coalgebraic modal logic. Propositional modal logic is equivalent to the bisimulation-invariant fragment of fixed point monadic second-order logic.[2] The fixed point higher-order modal logic is referred to as the modal μ-calculus, where $\mu(x)$ is an operator recording a least fixed point. Despite the non-transitivity of sorites phenomena – such that, on its epistemic interpretation, the subsequent invalidation of modal axiom 4 entails structural, higher-order epistemic indeterminacy – the modal μ-calculus provides a natural setting in which a least fixed point can be defined with regard to the states instantiated by non-deterministic modal automata. In virtue of recording iterations of particular states, the least fixed points witnessed by non-deterministic modal automata provide, then,

---

[2]Cf. Janin and Walukiewicz (1996).

an escape route from the conclusion that the invalidation of the KK principle provides an exhaustive and insuperable obstruction to self-knowledge. Rather, the least fixed points countenanced in the modal μ-calculus provide another conduit into subjects' knowledge to the effect that they know that a state has a determinate value. Thus, because of the fixed points definable in the modal μ-calculus, the non-transitivity of the similarity relation is yet consistent with necessary conditions on epistemic determinacy and self-knowledge, and the states at issue can be luminous to the subjects who instantiate them.

In the remainder of the essay, we introduce labeled transition systems, the modal μ-calculus, and non-deterministic Kripke (i.e., μ-) automata. We recount then the sorites paradox in the setting of the modal μ-calculus, and demonstrate how the existence of fixed points enables there to be iterative phenomena which ensure that – despite the invalidation of modal axiom 4 – iterations of mental states can be secured, and can thereby be luminous.

A labeled transition system is a tuple comprised of a set of worlds, M; a valuation, V, from M to its powerset, P(M); and a family of accessibility relations, R. So LTS = $\langle$M,V,R$\rangle$ (cf. Venema, 2012: 7). A Kripke coalgebra combines V and R into a Kripke functor, σ_R; i.e. the set of binary morphisms from M to P(M) (op. cit.: 7-8). Thus for an s$\in$M, $\sigma(s) := [\sigma_V(s), \sigma_R(s)]$ (op. cit.). Satisfaction for the system is defined inductively as follows: For a formula ϕ defined at a state, s, in M,

$\llbracket \phi \rrbracket^M = $ V(s) [3]
$\llbracket \neg\phi \rrbracket^M = $ S $-$ V(s)
$\llbracket \bot \rrbracket^M = \emptyset$
$\llbracket T \rrbracket^M = $ M
$\llbracket \phi \vee \psi \rrbracket^M = \llbracket \phi \rrbracket^M \cup \llbracket \psi \rrbracket^M$
$\llbracket \phi \wedge \psi \rrbracket^M = \llbracket \phi \rrbracket^M \cap \llbracket \psi \rrbracket^M$
$\llbracket \diamond_s \phi \rrbracket^M = \langle R_s \rangle \llbracket \phi \rrbracket^M$
$\llbracket \square_s \phi \rrbracket^M = [R_s] \llbracket \phi \rrbracket^M$, with
$\langle R_s \rangle (\phi) := \{s'\in$S $\mid R_s$[s'$] \cap \phi \neq \emptyset \}$ and
$[R_s](\phi) := \{$s'$\in$S $\mid R_s$[s'$] \subseteq \phi \}$ (9).

In propositional dynamic logic (PDL), $\langle \pi \rangle \phi$ abbreviates that some execution of a non-deterministic computable program entrains the information state contained in ϕ, where computability is here defined in accord with the Church-Turing thesis that a function is effectively computable if and only if

---

[3] Alternatively, M,s $\Vdash \phi$ if s$\in$V(ϕ) (9).

it is partial and recursive, as co-extensive with the class of λ-definable terms and the class of finite, discrete-state automata such as Turing machines (cf. Church, 1936; Turing, 1937). $[\pi]\phi$ abbreviates that all executions of a non-deterministic computable program entrains the information state contained in ϕ. Complex operations in propositional dynamic logics may then obtain (cf. Blackburn and van Benthem, 2007: 59-61). A Choice principle states that the union of π_1 and π_2 may be formed, such that the logic executes either π_1 or π_2. A Composition principle states that there is an operation $\pi_1;\pi_1$, such that the logic first executes π_1 and then executes π_2. An Iteration principle defines a program, π^*, where π^* entrains the execution of π a finite number of times. Finally, a Test principle defines a program, $\pi?$, where $\pi?$ can comprise the following algorithms: '$(\pi?;a) \cup (\neg\pi?;b)$', which states that if a program π obtains, then a obtains, else b obtains; 'a;$(\neg\pi?;a)^*;\pi?$', which states that the logic will repeat the execution, a, a finite number of times until the program π is tested; and '$(\pi?;a)^*;\neg\pi?$', which states that while a program π is being executed a finite number of times, do a (op. cit.: 59-60).

The modal μ-calculus is then defined as follows. Recall again the foregoing Iteration principle from PDL, $\langle\pi^*\rangle\phi$ (Venema, op. cit.: 25). In our Kripke colagebra, we thus have M,s $\Vdash$ $\langle\pi^*\rangle\phi$ $\iff$ $(\phi \vee \diamond_s\langle\pi^*\rangle\phi)$ (op. cit.). $\langle\pi^*\rangle\phi$ is thus said to be the *fixed point* for the equation, x $\iff$ $\phi \vee \diamond$x, where the value of the formula is a function of the value of x conditional on the constancy in value of ϕ (op. cit., 38). The smallest solution of the formula, x $\iff$ $\phi \vee \diamond$x, is written μ.x$\phi \vee \diamond$x (25). The value of the least fixed point is, finally, defined more specifically thus:

$\llbracket\mu.\text{x}\phi \vee \diamond\text{x}\rrbracket = V(\phi) \cup \langle\text{R}\rangle(\llbracket\mu.\text{x}\phi \vee \diamond\text{x}\rrbracket)$ (38).

A non-deterministic automaton is a tuple $\mathbb{A} = \langle\text{A}, \delta, \text{Acc}, \text{a}_I\rangle$, with A a finite set of states, a_I being the initial state of A; δ is a transition function s.t. δ: A $\rightarrow$ P(A); and Acc $\subseteq$ A is an acceptance condition which specifies admissible conditions on δ (60, 66). A Muller acceptance condition is defined as a subset of A, $\alpha \subseteq$ P(A), such that $\text{Acc}_\alpha := \{\text{s}\in\text{A}^\omega \mid \text{Inf(s)} \in \alpha\}$ (intuitively: the admissible states are the infinite states in a deployment of the transition function) (60). A Büchi condition is a subset $\beta \subseteq$ A, such that $\text{Acc}_\beta := \{\text{s}\in\text{A}^\omega \mid \text{Inf(s)} \cap \beta \neq \emptyset\}$ (intuitively: an operation of the transition function passes through the state s infinitely often) (op. cit.). Finally, a parity condition is defined via a mapping, Ω: A $\rightarrow \omega$, such that $\text{Acc}_\Omega := \{\text{s}\in\text{A}^\omega \mid \max\{\Omega(\text{s}) \mid \text{s} \in \text{Inf(s)}\}$ is an even number$\}$ (intuitively: Ω is the largest natural number occurring infinitely often in the sequence of states figuring as input to δ, and such that the automaton accepts a particular

infinite state iff Ω is even (op. cit.).

Let two Kripke models $\mathbb{A} = \langle A, a \rangle$ and $\mathbb{S} = \langle S, s \rangle$, be bisimilar if and only if there is is a non-empty binary relation, $Z \subseteq \mathbb{A} \times \mathbb{S}$, which is satisfied, if:

(i) For all $a \in \mathbb{A}$ and $s \in \mathbb{S}$, if aZs, then a and s satisfy the same proposition letters;

(ii) *The forth condition.* If aZs and $R_\triangle a, v_1 \ldots v_n$, then there are $v_1 \ldots v'_n$ in $\mathbb{S}$, s.t.

• for all i $(1 \leq i \leq n)$ $v_i Z v'_i$, and

• $R'_\triangle s, v'_1 \ldots v'_n$;

(iii) *The back condition.* If aZs and $R'_\triangle s, v'_1 \ldots v'_n$, then there are $v_1 \ldots v_n$ in $\mathbb{A}$, s.t.

• for all i $(1 \leq i \leq n)$ $v_i Z v'_i$ and

• $R_\triangle a, v_1 \ldots v_n$ (cf. Blackburn et al, 2001: 64-65).

Bisimulations may be redefined as *relation liftings.* We let, e.g., a Kripke functor, **K**, be such that there is a relation $\mathbf{K!} \subseteq \mathbf{K}(A) \times \mathbf{K}(A')$ (17). Let Z be a binary relation s.t. $Z \subseteq A \times A'$ and $P!Z \subseteq P(A) \times P(A')$, with

$P!Z := \{(X,X') \mid \forall x \in X \exists x' \in X' \text{ with } (x,x') \in Z \land \forall x' \in X' \exists x \in X \text{ with } (x,x') \in Z\}$ (op. cit.). Then, we can define the relation lifting, **K!**, as follows:

$\mathbf{K!} := \{[(\pi,X), (\pi',X')] \mid \pi = \pi' \text{ and } (X,X') \in P!Z\}$ (op. cit.).

Finally, given the Kripke functor, **K**, **K** can be defined as the μ-automaton, i.e., the tuple $\mathbb{A} = \langle A, \delta, \Omega, a_I \rangle$, with $a_I \in A$ defined again as the initial state in the set of states A; Ω defined once more as the foregoing parity acceptance condition; and δ defined as a mapping such that $\delta : A \to P_\exists(\mathbf{K}A)$, where the $\exists$ subscript indicates that $(a,s) \in A \times S \to \{(a',s) \in \mathbf{K}(A) \times S \mid a' \in \delta(a)\} = \Omega(a)$, and $(a',s) \in \mathbf{K}(A) \times S$ if and only if $\{Z \subseteq A \times S \mid [a', \sigma(s)] \in \mathbf{K!}(Z)\} = 0$ (93).

The philosophical significance of the foregoing can now be witnessed by defining the μ-automata on an alphabet; in particular, a non-transitive set comprising a bounded real-valued, ordered sequence of chromatic properties. Whereas Ω, in the above parity mapping, is identified with the largest even number occurring infinitely often in the alphabet over which the automaton is defined, the Muller acceptance condition would appear to be more suitable for a background language of real-valued, chromatic properties. The Kripke functor whose acceptance conditions are Muller permits us, subsequently, to define fixed points relative to arbitrary points comprising the non-transitive sequence. Thus, although the non-transitivity of the ordered sequence of

color hues belies modal axiom 4, such that one can know that a particular point in the sequence has a particular value although not know that one knows that the point satisfies that value, the perceived constancy of the chromatic values, ϕ, in the non-transitive set of colors nevertheless permits every sequential input state in the μ-automaton to define a fixed point. With $\delta^M_x(\text{x'}) := \langle \delta \rangle^{S[x \mapsto x']}$, the transition function can then satisfy the Muller condition relative to each point in the continuum, such that $\delta^M_x(\text{x'})$ iff $V(\phi \cup \langle \text{R} \rangle(\text{x'})$ iff $V(\phi) \cup \langle \text{R} \rangle([\![\mu.\text{x}\phi \vee \diamond \text{x}]\!])$ (38).

The epistemicist approach to vagueness relies, as noted, on the epistemic interpretation of the modal operator, such that the invalidation of transitivity and modal axiom 4 ($\Box \phi \rightarrow \Box \Box \phi$) can be interpreted as providing a barrier to a necessary condition on self-knowledge. Crucially, μ-automata can receive a similar epistemic interpretation. Thus, interpreting the μ-automaton's Kripke functors epistemically permits the *iterations* of the set of functions – as defined by the fixed points relative to the arbitrary points in the ordered continuum – to provide a principled means – distinct from the satisfaction of the KK principle – by which to account for the pertinent iterations of epistemic states unique to an agent's self-knowledge.

Chapter 8

Bayesian Apriority

This chapter aims to provide a precise formulation of the dependence relations between the apriori and aposteriori epistemic profiles of mental states. The significance of the result concerns the following. One delineation of the distinction between the target notions individuates the aposteriori by its evidential and enabling roles. An aposteriori truth satisfies the evidential role, if its epistemic import is one of either justification or knowledge. An aposteriori truth satisfies the enabling role, if the relation to which it stands to apriori truths is one of dependence, without entraining evidential justification thereby. An argument against the robustness of the distinction between the notions of the apriori and aposteriori can then proceed by arguing that the aposteriori stands in some intermediate relation to the apriori. If the aposteriori bears this intermediate relation to the apriori, then, the argument runs, 'there is no significant epistemological difference between' the two notions (cf. Williamson, 2013: 300).

In this chapter, I endeavor to reverse the foregoing argument. I argue that there is an alternative means of formally defining and understanding the foregoing terms, and that it is a theoretical virtue of the account that it is sufficient for capturing the intermediary conditions, between the enabling and evidential. I demonstrate how precisifying the multiple relations of dependence between the apriori and aposteriori: (i) can yield a better grasp of the distinct epistemic contributions of each epistemic property; and (ii) can further capture the properties of other notions pertinent to the resolution of the puzzles which concern the varieties of epistemic source.

In Section **2**, I provide definitions for the terms and proposals at issue, and I discuss the properties of their interaction. In Section **3**, I provide an in-

formal characterization of the framework, as well as of the desiderata that the model is intended to secure. In Section **4**, I provide the formal regimentation of the model characterized informally in **3**. In Section **5**, I demonstrate how the framework can be applied, in order to provide a resolution to a puzzle concerning the justification for immunity to error through misidentification. Finally, a second extension of the model is examined, in order to account for the precise interaction between epistemic modality and the semantic values of deontic modality. The signficance of the last extension is a precise account of the relations between apriority, explanation, and epistemic responsibility. Section **6** provides concluding remarks.

8.1   Definitions

Williamson provides the following provisional definitions for the terms, 'apriori', 'aposteriori', 'dependence', and 'experience' (op. cit.: 292-294).

The dependence relation holds (i) between the apriori and aposteriori profiles, on the one hand, and instances of experience types; and (ii) between the notions of the apriori and the aposteriori themselves.

With regard to the remit of (i): A means of generating an epistemic state is apriori if and only if it is independent of experience. A means of generating an epistemic state is aposteriori if and only if it depends on acquaintance with instances of an experience type.

With regard to the remit of (ii): An apriori mental state, A, can depend on an aposteriori mental state, M, if and only if M satisfies the enabling role. An aposteriori mental state enables an apriori mental state if and only if M individuates the concepts comprising A, but is otherwise silent with regard to the status of A as sufficiently comprising justification or knowledge (op. cit.).

Finally, an entity is an instance of an experience type, either if the entity is a token of 'outer' sense and comprises sensory phenomenal properties and perceptual intentional states, or if the entity is a token of 'inner' sense and comprises cognitive phenomenal properties unique to an agent's cognitive states of belief and introspection (op. cit.).

Williamson's taxonomy provides a compelling approach to the delineation of the notions of apriority and aposteriority. However, in the remainder of the chapter, I will endeavor to outline and develop an alternative, precise framework, which will aim to capture more of the defining theoretical con-

tours of each notion. I will argue, then, that a more fine-grained approach is necessary, in order to help illuminate the reasons for which a division between the apriori and aposteriori is theoretically robust.

8.2 An Informal Characterization of the Theory

Let the 1-intension of a sentence be a function from the sentence at an epistemically possible world to its semantic value. The 2-intension of a sentence is a function from the sentence at a metaphysically possible world to its semantic value. A sentence is super-rigid if and only if it is individuated relative to both the 1- and 2-parameters – i.e., its full 2D-intension – and is then true in all epistemically possible worlds and true in all metaphysically possible worlds. A sentence is conceivable if and only if it is epistemically possible, true at some epistemically possible world. A sentence is apriori if and only if it is epistemically necessary, true at all epistemically possible worlds if and only if it is inconceivable for the sentence to be false. A sentence is aposteriori if and only if it is either metaphysically possible, true at some metaphysically possible world, or metaphysically necessary, and thus true at all metaphysically possible worlds.

There are four crucial points at which the apriori can precisely depend upon the aposteriori, so defined.

(I) The first crux concerns the second-order implicit definitions for the first-order terms in epistemic and metaphysical logical space. Because they define the extensions of terms through the whole of epistemic and metaphysical logical space, second-order implicit definitions must themselves be epistemically necessary (i.e., apriori) and metaphysically necessary. To capture this idea, second-order implicit definitions can be understood as being modally factive apriori. A sentence is modally factive apriori if and only if it is *de iure* epistemically necessary and further witnessed by metaphysically possible worlds. Crucially, then, the condition precisifies a formal dependence relation between the apriority of concepts and the metaphysical properties which serve to individuate them. Working backward: At metaphysical worlds where the sentence is false, this would entrain the falsity of the sentence's 1-intension, and thereby compromise its apriori status.

(II) The second crux precisifies further the status of the modally fac-

tive apriority of second-order implicit definitions, by introducing subjective Bayesian confirmation measures and demonstrating the relevance thereof. The subjective probability operators which figure in the confirmation relations are themselves defined on epistemically possible worlds. However, the measures provide conditions for how second-order implicit definitions of first-order terms can relate to evidential confirmation.

(III) The third crux concerns the individuation-conditions on the intensions and propositions in epistemic modal space. Intensions are functions from epistemically possible worlds to their extensions. However, one problem with simply availing of this interpretation as it stands is that the extension of terms and formulas in the space of an agent's epistemic possibilities considered as actual has traditionally been taken to be a paradigm of narrow or internal content (cf. Chalmers, 2012). To avoid this result, one can argue that there are non-descriptive, externalist conditions on the intensions and propositions comprising epistemic modality. On the approach here advanced, the universe at issue is the cartesian product of (i) a Bayesian perceptual state space, and (ii) the space of states of information comprising epistemic modality.

(IV) The fourth crux concerns, finally, the status of hyperintensional operators, when interpreted as concerning the *epistemic* explanatory connection between truths. Beyond the subjective Bayesian confirmation measures, which again express the relation between second-order implicit definitions and evidence, grounding operators express the epistemically explanatory connection between epistemic norms and the confirmation measures for subjective probability. Grounding operators express, then, in virtue of what the subjective Bayesian confirmation measures, which codify the relation between evidence and second-order implicit definitions, are true epistemic norms.

The explanatory connection is epistemic, because the operator is defined on the space of epistemic possibilities, i.e. global states of information. Epistemic explanatory ground coincides, in part, with the notion of the 'relativized apriori', as advanced in Friedman (2001). Friedman (op. cit.) argues that scientific theories are relativized apriori, if there is a plurality of distinct sets of laws which figure as transcendentally 'presupposed', constitutive conditions in the models of the theory (as historically evinced by the advent of non-Euclidean geometry). Hyperintensional operators defined on epistemically possible worlds capture the explanatory dependence among truths in epistemic modal space. Epistemic hyperintensional dependence can thus be defined as a species of the relativized apriori, given the multitude of logics

and axioms one might endorse in constructing intended models which aim to satisfy abductive conditions on theoretical adequacy.

An iterated grounding operator explains in virtue of what a previous grounding operator is sufficiently explanatory. Once the epistemic explanatory bridge between, e.g., formal confirmation measures and epistemic norms has been forged, an inquiry arises as to in virtue of what the epistemic norms are valid. Thus, grounding operators at higher-order provide an explanation of the first-order explanatory connection between epistemic norms and confirmation measures. For example, higher-order grounding operators target the explanation with regard to which logic ought to govern the antecedent, explanatory connection between confirmation measures and the epistemic norms implicated by first-order grounding. Iterated grounding operators serve to explain, e.g., which assignment of logical values to credences are tautologous. The issue of which values in the inductive logical setting are tautologies would potentially yield a revisionary bearing on other probabilistic principles, such as Normality – which claims that certainty will map to 1 if the logic is classical, or .5 in three-valued paraconsistent and paracomplete logics – and on the Regularity principle – which claims that a sentence, ϕ, is epistemically possible only if it has a non-zero probability.

Suppose, then, that there is an epistemic explanatory connection, such that the validity of an epistemic norm depends upon one's choice of logic. Validity provides an interesting test case. Field (2008) draws, for example, on Curry's Paradox, in order to argue against identifying validity with necessary truth-preservation. The argument from Curry's Paradox is such that – by truth-introduction ['$\phi \rightarrow \mathrm{T}(\phi)$'] and truth-elimination [' $\mathrm{T}(\phi \rightarrow \phi$'] – one can derive the following.

$\phi \iff [\mathrm{T}(\phi) \rightarrow \bot]$
$\mathrm{T}(\phi) \iff [\mathrm{T}(\phi) \rightarrow \bot]$
$\mathrm{T}(\phi) \rightarrow [\mathrm{T}(\phi) \rightarrow \bot]$
$[\mathrm{T}(\phi) \wedge \mathrm{T}(\phi)] \rightarrow \bot$ (by importation)
$\mathrm{T}(\phi) \rightarrow \bot$
$[\mathrm{T}(\phi) \rightarrow \bot] \rightarrow \mathrm{T}(\phi)$
$\mathrm{T}(\phi)$
$\bot$

The paradox is intended to demonstrate that necessary truth-preservation entails contradiction. However, the argument need not be valid, if one retains truth-introduction and truth-elimination, and yet opts for a weaker logic. One can preserve truth-introduction and truth-elimination, by availing of

the strong Kleene valuation scheme, such that the value of ϕ is ungrounded; i.e. maps to .5.[1] One can then add an indeterminacy operator, such that it is not determinately true that ϕ and it is not determinately true that not ϕ; so, it is indeterminate whether ϕ. Field argues, in virtue of the foregoing, that validity ought not to be identified with necessary truth-preservation, and ought instead to be taken as primitive.

The grounding operator comes in types and with distinct interpretations. While the traditional interpretation has targeted the metaphysical interpretation of the explanatory grounding operator, the above proposal is that the operator can be interpreted as an epistemic explanatory connection. However, the two interpretations can interact. Thus, the validity of an epistemic norm might fully depend, for its explanation, on one's choice of logic. However, one's choice of logic might partly depend for its explanation on considerations from metaphysics. To complete the present example, suppose that the epistemic norm in question concerns validity. As its epistemic explanatory ground, grasp of the notion of validity will depend on one's choice of logic, e.g., K3 and indeterminacy. Beyond, though, its merits as a fix to the alethic paradoxes, in virtue of what should one adopt K3? The inquiry targets an iteration of the grounding operator, concerned now with the metaphysical explanation for adopting the logic. Here is one such metaphysical consideration: Suppose that one distinguishes between fundamental and derivative metaphysical states of affairs. The fundamental states of affairs might concern the entities located in $3n$-dimensional spacetime, such as the field represented by the wavefunction. The derivative states of affairs might concern the entities located in lower, 3-dimensional spacetime, such as the clouds. One can then say that the fundamental states of affairs take the classical values, [0,1]. In order to capture the priority of the fundamental to the derivative, one can argue that the derivative states of affairs are governed by K3 and indeterminacy, such that – while fundamental states are always either true or false – it is not determinate that a derivative state of affairs obtains, and it is not determinate that a derivative state of affairs does not obtain.[2]

---

[1]Cf. Kleene (1952).

[2]Russell (2013: 12-14) advances a 'factuality' operator, which nearly converges with the metaphysical application of K3 and indeterminacy that I have proposed here. One major difference between the approaches is that Russell does not provide a definition of 'It is factual that', preferring to take the notion as primitive. He works, further, in a two-valued setting. The application of the factuality operator is intended to capture both the

In this manner, there can be an epistemic explanatory connection between the status of an epistemic norm such as validity and the status of one's choice of logic. As evinced by the foregoing, there is, furthermore, a metaphysical ground for the epistemic explanatory connection: One's choice of logic can be partly grounded by considerations of metaphysical dependence.[3]

In the following section, I will examine how the foregoing conditions can be formally expressed. Section **5** examines one of two applications of the model. The first extension of the model is to the puzzle of first-person immunity, and exploits the apparent tension between apriority and aposteriority. I demonstrate how – by availing of the present framework – a resolution to the puzzle can be forthcoming. The second extension of the model is to an account of how the space of epistemic modality and the space of deontic modality on various interpretations can precisely interact.

The significance of the framework is that it demonstrates how – rather than undermining the distinction between the apriori and the aposteriori – the multiple, fine-grained relations between evidential and enabling conditions are crucial: (i) for a resolution of the various puzzles that derive from

notion of the qualitative and the types of objectivity: Factual propositions are qualitative, and thus taken to supervene on perfectly natural properties. Thus, propositions mapping to 1 can be 'factually defective' and satisfy only deflationary conditions on their truth, whereas there are also propositions mapping to 1 whose factuality is in some sense more realistic. Russell argues, then, that there are possibilities that are not possible worlds, because possibilities can be factual or non-factual depending on whether they supervene on the qualitative, natural properties of the worlds in which they exist. By contrast, the proposal that I have outlined is that K3-plus-indeterminacy is sufficient for capturing the metaphysical distinction between fundamental states of affairs and derivative states of affairs.

[3]Another example of how one's choice of logic might depend upon considerations of metaphysics is Fine's (2013) truthmaker semantics for intuitionistic logic. Considerations of how states and fusions thereof can provide an exact verification for derivative states might vindicate, for this purpose, the adoption of intuitionist negation. A third example, which is unique in targeting sub-structural logic, can be found in Cotnoir's (2015) work on Abelian intensional mereology. A mereological group is intensional if and only if two groups with the same parts are distinct. Thus, in some sense, the parts can figure more than once in the whole. This would be a violation of mereological 'Idempotence', which states that 'If x is a part of y, then the sum of x and y is identical to y'. There might, e.g., be distinct, admissible ways of counting the parts of a whole, such that the parts figure more than once therein. Further, if the parts of a whole can be multiply located at a time, then this might be an instance in which the parts of the whole can be witnessed more than once. So, if mereological groups are intensional, then one's logic ought to invalidate Idempotence.

the interaction between the apriori and aposteriori; and thus (ii) for witnessing the epistemological significance of the distinction between the two notions.

8.3   The Formal Regimentation

This section provides a formal account of the informal theory proffered in the last section. I provide informal elucidations of the formal clauses, subsequent to defining them precisely. The informal definitions adumbrated in the previous section are further intended to provide an intuititive characterization of the formalism.

8.3.1   Epistemic Modality

- **1-Intension:**

  $\mathtt{pri}(x) = \lambda v.[\![x]\!]^{v,v}$

  (the two parameters relative to which x obtains its value are epistemically possible worlds).

- **2-Intension:**

  $\mathtt{sec}_{v_@}(x) = \lambda w.[\![x]\!]^{v_@,w}$

  (the value of x is relative to a parameter for the space of epistemically possible worlds, one of which is designated as actual. The value of x relative to the first parameter determines the value of x relative to the second parameter for the space of metaphysically possible worlds).

The foregoing conceivable intensions are individuated by mental images, where the latter are a species of phenomenal property type. Phenomenal properties are defined on a state space, Ω, which represents the possibilities concerning, e.g., the direction of a light source.

- **Phenomenal Properties**

  Comprehension:

  $\mathtt{Comp} = \lambda\alpha\forall x.\alpha(x) \Longleftrightarrow A$

  with α not free in A.

Phenomenal properties are picked out by the concepts thereof.

- **Phenomenal Concepts**

  Γ denotes a 2nd-order concept

  $\Gamma[\lambda\alpha_1.[\alpha_1(v)]] = \Gamma[\lambda\alpha_2.[\alpha_2(v)]]$ iff μ

  where v is bound by a universal quantifer ranging over a domain of individuals, $D = \{x_1 \ldots x_n\}$; α_i denotes a member of the domain of phenomenal properties; and μ is a condition in which neither x_i nor α_i occur free.

Concepts of phenomenal character can be defined by availing of the epistemic interpretation of two-dimensional intensional semantics. Phenomenal concepts are individuated by double-indexed intensions, where the intensions are *de iure* epistemically and subjunctively rigid, s.t.

- **Super-rigidity (2D-Intension):**

  $[\![\Gamma]\!]^{c,w} = 1 \iff \forall c',w'[\![\Gamma]\!]^{c',w'} = 1$

The super-rigidity condition for phenomenal concepts is intended to secure the idea that there is no gap between knowledge or the appearance of a certain phenomenal property – e.g., pain – and the metaphysical instantiation thereof.

Phenomenal properties do not merely float free of reality, possessed by no-one and further unencumbered by empirically tractable instantiation-conditions. The best candidate for a necessary condition on the instantiation of phenomenal properties on Bayesian perceptual state space is the operation of some kinds of attention. The kinds of attention include exogenous attention; endogenous attention; spatial-based attention; object-based attention; and property-based attention. If the grain of attention is coarser than the grain of conscious object-based perception, then object-based attention cannot figure as a necessary condition (cf. Block, 2013). Endogenous attention might, furthermore, fail to be a necessary condition, if the fundamental neural correlates of consciousness are ultimately localized to the operations of attention and vectorwave activity restricted to occipito-thalamic pathways (cf. Tsuchiya et al, 2013). Exogenous, spatial-based, and property-based attention are thus the remaining candidates for being necessary conditions on the instantiation of phenomenal properties. For the purposes of this note,

the necessary functional properties of the relevant kinds of attention can be targeted at the neurofunctional, rather than psychofunctional, level. They can be captured, e.g., by the 'Normalization Model', according to which 'responses of neurons are divided by a common factor that typically includes the summed activity of a pool of neurons' (Carandini and Heeger, 2011: 51. See also Reynolds and Heeger, 2009; and Chirimuuta, 2014):

- **Normalized Attention**

$$\bar{E}_i(n) = \frac{E_i(n)}{s^2 + \sum_i E_i(n)}^4$$

A random variable maps the possibilities in Ω to the values in the [0,1] interval. The random variable is interpreted as a probability density. The probability density is governed by the Kolmogorov axioms: namely, normality (which states that the probability of a tautology maps to 1); non-negativity (which states that the probabilty operator must take a non-negative value), and additivity (which states that for all disjoint probability densities, the probability of their union is identical to the probability of the first density added to the probability of the second). The phenomenal properties defined on Bayesian perceptual state space provide a non-descriptive and potentially externalist condition on the individuation of the intensions defined on epistemic worlds. The individuation can be codified by taking the cartesian product of Ω and epistemic modal space, V.

- **Normality**

  $Pr(\mathrm{T}) = 1$

- **Non-negativity**

  $Pr(\phi) \geq 0$

- **Additivity**

  If ϕ and ψ are disjoint, then,

  $Pr(\phi \cup \psi) = Pr(\phi) + Pr(\psi)$

- **Epistemic Necessity (Apriority):**

  $[\![\blacksquare\phi]\!]^{v,w} = 1 \iff \forall v'[\![\phi]\!]^{v,v'} = 1$

  (ϕ is true at all points in epistemic modal space).

---

[4]$E_i(\mathrm{n})$ denotes the firing-rate of a neural cell at certain oscillations, and s is a constant which encodes response saturation, i.e. the strength of sensory inputs as encoded by $E_i(\mathrm{n})$.

- **Epistemic Possibility**

  $[\![\diamond_v \phi]\!]^{i_v} \neq \emptyset \iff [\![Pr\phi]\!]^{i_v} \neq \emptyset \land >.5$, else $\langle \emptyset, Pr_{i_v}(\phi \mid \emptyset) \rangle$, where i_v designates an agent's state of information in a context.

  (ϕ might be true if and only if its value is not null and it is greater than .5).

- **Metaphysical (2-)Necessity:**

  $[\![\Box\phi]\!]^{v,w} = 1 \iff \forall w' [\![\phi]\!]^{v,w'} = 1$

  (ϕ is true at all points in metaphysical modal space).

- **Super-rigidity (2D-Intension):**

  $\mathtt{mat}(\phi) = [\![\phi]\!]^{v,w} = 1 \iff [\![\phi]\!]^{v',w'} = 1$

  (the intension of ϕ is rigid in all points in epistemic and metaphysical modal space).

- **Modally Factive Apriority:**

  $\blacksquare!(\phi) = \forall v' [\![\phi]\!]^{v',v'} = 1 \land \exists w' [\![\phi]\!]^{w,w'} = 1$

  (ϕ is true at all points in epistemic modal space, and is further true at one metaphysically possible world).

8.3.2   Multidimensional-Intensional Implicit Definition

- $\blacksquare!['\Gamma(x) = \Gamma(y)'$ iff $\bullet]$,

  where $\bullet$ denotes the set of conditions jointly necessary and sufficient for the identity and distinctness of Γ.

8.3.3   Evidential Confirmation

Let Φ denote the second-order implicit definition for the concept, Γ.

Let R denote the set of Bayesian confirmation relations between experience types, E, and Φ.

Then, $R = \{[\mathrm{Pr}(\Phi|E) - \mathrm{Pr}(\Phi)] \lor [\log[\mathrm{Pr}(\Phi|E)/\mathrm{Pr}(\Phi)]] \lor [\log[\mathrm{Pr}(E|\Phi)/\mathrm{Pr}(E|\neg\Phi)] \lor \ldots\}$

8.3.4 Grounding Structure

- G(`Epstm`, R)

An operator has a *ground-theoretic* interpretation, only if the profile induced by the interpretation concerns the *explanatory* connection between an antecedent set of truths and the relevant consequent.

We follow here the presentation in Fine (2012a, 2012b). Let a grounding operator be *weak* if and only if it induces reflexive grounding. A grounding operator is *strict* if and only if it is not weak. A grounding operator is *full* if and only if it uniquely provides the explanatory ground for a truth. A grounding operator is *part* if and only if a truth provides part of the explanation for a larger set of truths in which it figures.

Let Ψ denote the set of epistemic norms. Then:

- $\Psi < R$

  if and only if Ψ is a *strict full* ground for R

- $\Psi \leq R$

  if and only if Ψ is a *weak full* ground for R

- $\Psi \prec R$

  if and only if Ψ is a *strict part* ground for R

- $\Psi \preceq R$

  if and only if Ψ is a *weak part* ground for R

- $\Psi \preceq R \land \neg(R \preceq \Psi)$

  if and only if Ψ is a *strict partial* ground for R

- $\Psi \prec^* R$

  if and only if $\Psi_1, \ldots, \Psi_n \leq R$,

  if and only if Ψ is a *partial strict* ground for R;

- $\Psi \prec' R$

  if and only if $[\Psi \prec^* \mu \land \mu \preceq R]$,

  if and only if Ψ is a *part strict* ground for R.

114

8.3.5   Iterated Ground

- $G_3[G_2[G_1(\Psi,\ R)]] \iff G_2(\text{L-truth},\ \Psi)$ and $G_3(\bullet,\ \text{L-truth})$,

   where the interpretation of G_1 and G_2 concerns, as above, the epistemic explanatory connection between the truths at issue, the interpretation of G_3 concerns the metaphysical explanatory connection, and $\bullet$ designates members of the set of metaphysical truths.

8.3.6   Apriority and Abduction

- $\blacksquare\phi \prec \blacksquare\psi$,

   where $\phi := $ 'A'; and $\psi := $ 'A $\wedge$ B'

   (suppose that A is apriori and A $\wedge$ B is apriori. Then, A provides a strict part explanatory ground for the conjunction of A and B).

The foregoing provides the formal clauses for the model that was informally characterized in Section **3** above. In the following sections, I endeavor to demonstrate the philosophical significance of the model, by exploring extensions of the framework, both to the conditions on immunity to error through misidentification, and to the precise interaction between epistemic and deontic modalities. The puzzles exploit a more coarse-grained approach to the dependence relations between apriority and aposteriority. I examine, in turn, how the more fine-grained approach developed here is able to provide a resolution to each of the puzzles.

8.4   The First-Person Concept

8.4.1   The Empirical and Metaphysical Selves

In this section, I examine an extension of the present model to the resolution of a puzzle – immunity to error through misidentification – which hinges on the contrasting etiology and epistemic profiles of intentional states.

Fine (2005) distinguishes between two properties of subjectivity (312). The first is the property of being a 'metaphysical self'. The second is the property of being an 'empirical self'. A subject instantiates the property of being an empirical self if and only if the subject instantiates empirical properties (313). A subject instantiates the property of being a metaphysical self

if and only if a set of first-personal facts is indexed to that subject (314). The empirical properties that are instantiated by the subject, and which individuate their property of being an empirical self, are constitutive of the contents individuating the property of being a metaphysical self (317). Thoughts in which the first-person concept figures can thus be provisionally captured as follows: The first-person concept has as its extension the property of being a metaphysical self, if and only if, and because, (i) the concept has as its extension the set of first-personal facts that are indexed to a particular individual, and (ii) the individual instantiates the empirical properties comprising the set of first-personal facts. Thus, in the thought, 'I am NN', the first-person concept picks out the property of being the 'subject at which the egocentric facts obtain' (316). Fine avers that, in a proper set of egocentric facts, thoughts in which the first-person concept figures have the effect of 'restricting egocentric reality to the experience of experiences that only [a unique individual] can have' (318).

We can be more precise. The empirical self is the set of experiences: (i) instantiated by an individual, and (ii) comprising the egocentric facts.

- **ES**

  $\lambda\eta\exists\iota x.(\{w:\ i_1,\ \ldots,\ i_n\} = \eta)](x)$

The metaphysical self is the property of being the subject at which the egocentric facts obtain, the second-order property of instantiating the first-order property, η:

- **MS**

  $\lambda M\lambda\eta.[\exists x[M(\eta)]](x).$

I will take the set of egocentric facts to refer to an individual's total phenomenal evidence. Phenomenal evidence is comprised of phenomenal properties. Given the conditions on phenomenal properties that were specified in Section **4**, the first-person concept can, finally, be identified as follows:

- **The First-Person Concept**

  $I(x) = I(y)$ iff $\lambda\eta\lambda M\exists\iota x.(\{w:\ i_1,\ \ldots,\ i_n\} = \eta)](x) \wedge [M(\eta)](x)$

8.4.2 Super-Rigidity and Immunity

Evans (1985: 307) defines the first-person concept as follows:

- $\lambda \gamma . R_I (\gamma,\, x,\, t)$,

if

- `Identification`:

  $\gamma = x$.

γ denotes a primitive mode of presentation; x denotes the objctive profile of an individual; and t denotes an instant of time.

Cases in which the satisfaction conditions on `Identification` are met are standardly referred to as cases of 'immunity to error through misidentification' (cf. Shoemaker, 1968; Evans, 1982: 215).[5] Among the satisfaction conditions that have been proffered are:

(*) a species of default, apriori entitlement to presuppose `Identification` (cf. Wright, 2012); and
(**) the contrasting proposal that Evans' primitive mode of presentation is reducible to proprioceptive and other sensory properties, which are unique to the individual who instantiates them (cf. Recanati, 2007: Ch. 6).

The formal and philosophical considerations adduced in Section **4** demonstrate that both the defining contours of (*) and (**) are correct. Only one, unique individual can satisfy the instantiation-conditions on the set of sensory phenomenal properties, which individuate that unique individual's second-order property of being a metaphysical self.

With regard to (*):
The *apriority* of the entitlement to presuppose `Identification` is secured by both the definition of apriority as epistemic necessity, and the modal profile of phenomenal intentional properties. Phenomenal intentional states are *de iure* epistemically rigid. Phenomenal concepts thus pick out the same unique set of properties in all epistemically possible worlds.

---

[5]Cf. Pryor (1999), for a generalization of Evans' principle and the `Identification` component to all singular terms and the objects that they serve to designate.

With regard to (**):
The epistemic rigidity of the constitutive properties falling under the first-person concept is consistent with the *evidential* satisfaction conditions on `Identification`, in virtue of the following. The precise mechanism by which the set of phenomenal properties uniquely belongs to an individual, and comprises the second-order property of being a unique metaphysical self, just is the modal profile of phenomenal properties. Phenomenal intentional states are subjunctively rigid. Phenomenal concepts thus denote the same unique set of properties at all metaphysically possible worlds.

Hence, the first-person concept picks out the unique individual to whom a unique set of constitutively *de iure* epistemically and metaphysically rigid, phenomenal intentional properties can be attributed. The modal profile of phenomenal properties – i.e., the property of being *de iure* epistemically and metaphysically rigid – is not reducible to the empirical instantation-conditions of the latter. Rather, the empirical properties availed of in the foregoing possess an implicitly modal profile. The significance of examining the relevant empirical instantiation-conditions – captured, e.g., by the precise models of attentional mechanisms and the perceptual state space on which phenomenal properties are defined – is a naturalistically adequate explanation, both of the properties of empirical and metaphysical selfhood, and of the manner in which an individual's personal identity can be identical with their epistemically modalized yet modally factive, total phenomenal evidence.

8.4.3   Immunity and Indeterminacy

In the foregoing, the points in the Bayesian state space on which phenomenal intentional properties are defined was assumed to take classical maximal values. However, there are cases in which the truth of the `Identification` principle exhibits a species of indeterminacy, both on its epistemic and metaphysical interpretations.

One example of indeterminacy with regard to the extension of the first-person concept might be Anscombe's (1975) sensory deprivation scenario. Anscombe writes:

'[I]magine that I get into a state of 'sensory deprivation'. Sight is cut off, and I am locally anaesthetized everywhere, perhaps floated in a tank of tepid water; I am unable to speak, or to touch any part of my body with any other. Now I tell myself 'I won't let this happen again!' If the object meant

by 'I' is this body, this human being, then in these circumstances it won't be present to my senses; and how else can it be 'present to' me? But have I lost what I mean by 'I'? Is that not present to me? Am I reduced to, as it were, 'referring in absence' (op. cit., 57)?

Anscombe's objection rests on the assumption that thoughts in which the first-person concept figures are entertainable, only if the concept picks out the subject of egocentric facts (i.e., a set of phenomenal properties) at the instant of time that the thought is entertained. The present proposal rejects such an assumption. The identity-conditions on the first-person concept are the instantiation, by an individual, of a set of phenomenal properties – i.e., the empirical self – and the second-order predication of that set of properties to the individual – i.e. the metaphysical self. Immunity to error through misidentification is justified because phenomenal properties have an implicitly modal profile. Knowledge and the concepts of phenomenal character have the property of being *de iure* epistemically rigid. Knowledge and concepts of phenomenal properties are also *de iure* subjunctively rigid. Thus, there is no gap between the appearance of the phenomenal properties which individuate the first-person concept and their actual metaphysical instantiation. The extension of the concept is *de iure* epistemically and subjunctively rigid, if and only if the individuation-conditions on the concept have, at some time in the past, been satisfied. Thus, against Anscombe's objection, the subsequent deployment of the concept in thought does not require that the individuation-conditions on the concept are perpetually being satisfied.

Another example of the interaction between immunity and indeterminacy targets the phenomenon of indeterminate survival.[6] van Inwagen (op. cit.) presents the phenomenon via the following example:

'Suppose that a person, Alpha, enters a certain infernal philosophical engine called the [Wardrobe]. Suppose that a person later emerges from the [Wardrobe] and we immediately name him 'Omega'. Is Alpha Omega? [...] Alpha has entered and Omega has left. It is, therefore, not definitely true or definitely false that Alpha is Omega' (op. cit.: 243-244).

How does the *de iure* epistemic and subjunctive rigidity of the first-person concept interact with the indeterminacy of the identity of Alpha with Omega? Suppose that the indeterminacy at issue targets `Identification`, i.e. the relation between the primitive first-person mode of presentation and the ob-

---

[6]The locus classicus is van Inwagen (1990). Cf. Williams (2014a; 2014b), and Moss (2015), for further discussion.

jective profile of the individual. Thus, it is not determinate that Alpha's first-person mode of presentation will uniquely pick out Alpha's objective profile, and it is not determinate that it is not the case that Omega's first-person mode of presentation will uniquely pick out Omega's objective profile. In virtue of the indeterminacy that is dramatized by van Inwagen's example, the identification component would appear to satisfy the strong Kleene valuation scheme, such that it takes the value 'neither true nor false', governed by a determinacy operator.

Suppose that determinacy tracks fundamentality. As argued in Section **3**, one vindication of the strong Kleene valuation scheme as the semantics for metaphysics is that it can be deployed to secure the distinction between fundamental rather than derivative states of affairs. The fundamental states of affairs will receive the classical values, [0,1]. By contrast, the derivative states of affairs will receive the value, .5, corresponding to the alethic status 'neither true nor false', s.t. that it is not determinately true that the state of affairs obtains and not determinately true that the state of affairs does not obtain.

Suppose that the first-person concept picks out the subject of egocentric facts, i.e. a unique set of *de iure* epistemically and subjunctively rigid, phenomenal intentional states. How would the foregoing approach interact with the Wardrobe scenario?

If one believes that phenomenal property types are metaphysically fundamental, yet stand in some relation to physical truths – the position of the property dualist – then the value of `Identification` will be either true or false. This can be accommodated, if one were to argue that the indeterminacy of the Wardrobe scenario consists in the non-rigidity of the extension of the concept. Either the non-rigidity affects the result, at the level of metaphysically possible worlds, with regard to whether Alpha is Omega; or the non-rigidity affects the result, at the level of epistemically possible worlds, with regard to whether Alpha is Omega.

However, under what conditions can an epistemically and subjunctively rigid concept – e.g., the first-person concept – cease to be so? Barring neighborhood functions, epistemic and metaphysical necessity are too uncompromising to accommodate variations in their instantiation. If this is correct, then the indeterminacy at issue will have to be secured by another means. Thus, even if facts about phenomenal intentional properties are just as fundamental as facts about fundamental physical entities, the Wardrobe scenario demonstrates that – in some way – the indeterminacy of the identity of in-

dividuals must be secured.

If one takes phenomenal intentional properties to be fundamental, then the indeterminacy evinced by the Wardrobe scenario might have to consist in the failure of axiom 4 (on its epistemic interpretation) for the 1-intension of phenomenal concepts, which would thereby entrain that the rigidity of the 1-intension for the first-person concept is not itself known.[7] Axiom 4 states that '$\Box \phi \rightarrow \Box\Box\phi$'. Thus, on this approach, indeterminacy will be witnessed by the failure, in the Wardrobe scenario, of second-order knowledge of the first-order epistemic-rigidity of the first-person concept.

If one does not take phenomenal intentional properties to be fundamental, then propositions about the properties will be assigned the value .5, as yielded by the strong Kleene valuation, [0,.5,1]. The indeterminacy evinced by the Wardrobe scenario might, then, have to consist in the fact that phenomenal properties are derivative metaphysical entities, for which it is not determinately the case that states of affairs with regard to them obtain, and it is not determinately the case that states of affairs with regard to them do not obtain. Thus, the properties of consciousness, such as colors and pains, have as indeterminate a metaphysical profile as do other entities that are not metaphysically fundamental – e.g., the clouds – by contrast to the metaphysically fundamental entities that are postulated by fundamental physics. Thus, on this approach, indeterminacy will be witnessed by the extensional invariance of the values assigned to propositions which specify the extensions of phenomenal concepts in metaphysically possible worlds; namely, the third value corresponding to the status that it is not determinately the case that the concept picks out the property and not determinately not the case that the concept picks out the property.

8.5 Epistemic Modality and Deontic Modality

In this final section, I will examine the interaction between the semantic values of epistemic modal operators specified in the foregoing, and the varieties of deontic modal operators. This extension of the present framework is of crucial relevance for the following reasons. Work on the core of issues

---

[7]For the foundations of the epistemic approach to indeterminacy, see Sorensen (1988), and Williamson (1990/2013; 1994; forthcoming).

surrounding the notions of epistemic responsibility and the requirements of rationality, have drawn on both the resources of epistemic and deontic modality, without clearly delineating the properties of each notion or the properties of their interaction. The issue on which I will here focus concerns the justification concerning the correctness of belief-forming methods, in virtue of practical considerations which take the form of obligations on an agent to engage in certain explanatory tasks (cf. Enoch and Schechter, 2008).

The semantic values for the varieties of deontic modality are as follows:

A *subjective deontic modal* (taking scope over distinct atoms) maps to 1 – relative to context (agent, time), c, and index, i, on which a state space (of states of information) with preference-ranking and probabilistic structure is built:

$\llbracket \mathsf{O}_S(\text{x,y}) \rrbracket^{[c=(a,t)];id} = 1$, where

id $= \langle V, Pr, \text{Pref} \rangle$, s.t. $\{v \in V: i_1 \ldots i_n\}$; Pr is the probability density for the random variable mapping i_n to the [0,1] interval; and Pref is a strong ordering, s.t. for any i_x, i_y, and i_z, Pref is irreflexive, anti-symmetric, and transitive.

An *objective deontic modal* (taking scope over distinct atoms) maps to 1 – relative to context (agent, time), c, and index, i, on which a state space (of facts) with grounding structure is built:

$\llbracket \mathsf{O}_O(\text{x,y}) \rrbracket^{c=(a,t);id} = 1$, where

id $= \langle W, G \rangle$, s.t. $\{w \in W: G(w_x, w_y)\}$

In the deliberative setting – concerning whether a particular agent ought to ϕ – subjective deontic modals take explanatory priority to objective deontic modals. Subjective deontic modals concern, then, the target rational projects, engagement in which is a practical obligation, where the practical obligation codified by the deontic modal provides default epistemic justification for the correctness of the methods at issue. These projects include explaining one's surrounding phenomena, either by inference to the best explanation or via Bayesian confirmation, and the project of conditional presupposition with regard to future-planning (cf. Enoch and Schechter, op. cit.: 558).

In the evaluative setting – unrestricted to the states of information possessed by particular agents; and so concerning, e.g., how slavery ought to be forbidden, and how there ought to be unrestricted peace in 3- and 3*n*-dimensional spacetime – objective deontic modals take explanatory priority

to subjective deontic modals.[8] The explanatory priority between subjective and objective deontic modals in the evaluative and deliberative settings can, furthermore, be codified, by availing of the grounding structure for explanatory dependencies as detailed in Section **4**, above.

Thus, on the present proposal, epistemic possibility plays a constitutive role in subjective deontic modality, because it constitutes the space of states of information on the index in the array of intensional parameters, relative to which the subjective deontic modal maps to 1. Epistemic justification by default is a species of the apriori; and apriority is here represented by epistemic necessity.[9] Deontic modalities can provide default epistemic justification with regard to the correctness of belief-forming methods if and only if what the agent ought to do is governed by the modal system KT. $\mathsf{O}(\phi)$ states that 'It ought to be the case that ϕ', and $\mathsf{P}(\phi)$ states that 'It is permitted that it be case that ϕ'. Axiom K states, then, that '$\mathsf{O}[\phi \rightarrow \psi] \rightarrow [\mathsf{O}(\phi) \rightarrow \mathsf{O}(\psi)]$'. Axiom T states that '$\mathsf{O}(\phi) \rightarrow \phi$'. T can be understood as regimenting the thought, familiar from Leibniz (1714/1998: 53-60), that 'ought implies is'. Axiom T has appeared problematic in the setting of deontic modality. One reason adducing in favor of the avoidance of endorsing T in the deontic setting is that the actual prevalence of abhorrent cruelty and criminality

---

[8]See Schroeder (2011), for the distinction between evaluative and deliberative settings for the interpretation of the deontic modal. Barcan (1966) proffers a distinction between evaluative and prescriptive interpretations for the deontic modal, where the latter interpretation motivates the eliminability of iterated deontic operators; and thus presumably axiom 4, '$\mathsf{O}\phi \rightarrow \mathsf{O}\mathsf{O}\phi$'.

[9]The present examination targets thoughts expressing states of epistemic modality, rather than update-conditions on sentences in natural language semantics. Yablo (2011) proffers an approach to the interaction between epistemic and deontic modality, on the model of his extrapolation-based semantics for intentional states. [Cf. Yablo, 2014, for the foundations of his approach.] Yablo motivates the extension of his semantics for intentionality to epistemic and deontic modals, by targeting several issues for their interaction in the setting of an update semantics for natural language. By focusing on the semantic values of epistemic and deontic modal *judgments* rather than context-change potentials for sentences (hostage to the ambiguity and exigency of linguistic data), the present approach circumvents the linguistic concerns adducing in favor of Yablo's alternative proposal.

Extensions of the present model to decision theory can also be secured, by adding further indices on which states of spaces of acts and of outcomes can be defined. Some loci classici for the interaction between deontic modality and decision theory include Thomason (1981); Horty and Belnap (1995); Kolodny and MacFarlane (2010); Dowell (2013); and Cariani, Kaufmann, and Kaufmann (2013). For an approach to deontic modality which targets the evaluative interpretation of the modal, within an update semantics ordered by a global preference relation, see Starr (forthcoming).

should not be entailed by what ought to be the case. Axiom D, according to which '$O(\phi) \rightarrow P(\phi)$' would appear to be more innocuous: 'If it ought to be the case that ϕ, then it is permissible that ϕ' is occasionally true; e.g., that an agent ought to turn the other cheek is consistent with the permissibility of so acting, whereas that an agent ought to tear the fascist down might not be consistent with the permissibility of so acting.

However, a restricted version of axiom T is precisely what is needed, in order to complete the present account of how the interaction of epistemic and deontic modality is of relevance to an agent's epistemic responsibility, with regard to the justification for the correctness of their belief-forming methods. The restriction is formally induced by augmenting the model, M = $\langle$V, R, g$\rangle$ – where V denotes the space of epistemically possible worlds – with a neighborhood frame, $\langle$F,N$\rangle$ – where F is a subset of V, and N is a neighborhood function. Intuitively, N maps a subset of V to a particular world in V. Thus, '$O(\phi)$' in V if and only if 'N: $\phi \rightarrow$ v$\in$V', such that ϕ is necessarily true in the neighborhood structure, v, although not in the remainder of the full set of worlds comprising V.

An agent is therefore justified by default in pursuing a belief-forming method, because they are practically rationally obligated so to act. The agent is practially rationally obligated so to act, because pursuing the rational projects of explanatory and deliberative inquiry and self-evaluation is constitutive of being a reflective agent. (This is an instance of restricted T.) Thus, an agent possesses practical obligations with regard to the pursuit of belief-forming methods if and only if, and because, the agent does pursue projects of reason-governed inquiry. The agent does pursue the foregoing belief-forming methods, in virtue of their states of information and the facts comprising the cognitive psychological basis thereof.

A virtue of the present approach is its generality. The approach is general enough to accommodate several choice points, with regard to constraints on the array of intensional parameters for the values of the deontic and epistemic modals. These include, inter alia: (i) the normative weight of sociological properties in the setting of subjective deontic modality; (ii) restrictions of the axioms (such as T) in the deontic interpretation, in order to avoid the counterintuitive consequences entrained by their unrestricted versions; (iii) varying accounts of the dependence of ought-judgments on other value-theoretic considerations, such as e.g. the utility of certain acts or rules; (iv) either internalist or externalist conditions on the accessibility of information encoded by the state space of epistemic possibilities, relative to which the

124

values of subjective deontic modals are specified; and (v) the ontological dependencies between the space of facts, comprising the index relative to which the semantic values of objective deontic modality are specified.

8.6   Concluding Remarks

In this chapter, I have endeavored to delineate the precise contours of the various dependence relations – beyond the enabling and evidential – which hold between an intensional understanding of apriority and the notions of the aposteriori. The formally precise convergences between the apriori and aposteriori were intended to secure three, theoretical aims. First, the convergences demonstrate the ineliminable import of each notion. Because both formal measures of evidential confirmation and epistemically explanatory connections are definable on epistemically possible worlds, they provide the resources for explaining the manner in which the modally factive apriority of implicit definitions can depend upon evidential properties unique to the domain of the aposteriori. Second, the precise convergence yet ineliminable import of the apriori and aposteriori provide the resources necessary for an account of the phenomenon of conceptual fit, where the notion of conceptual fit arguably figures in the resolution of the puzzles concerning the contrasting etiology and epistemic profile of intentional states; e.g., the puzzle of first-person immunity. Third, I demonstrated the manner in which the proposal developed in the foregoing provides the resources necessary for a formally precise account, concerning the interaction between epistemic possibility and the distinct interpretations of deontic modal operators in the evaluative and deliberative settings. The philosophical significance of the interaction between the semantic values of epistemic and deontic modal operators was then examined, and was shown to explain in part the justification for the correctness of the belief-forming methods constitutive of reflective inquiry.

Rather than undermining the epistemological significance of the apriori and aposteriori divide, the multiple, formal points of intersection between the apriori and aposteriori are crucial (i) for a resolution of the various puzzles that derive from the interaction between the two notions, and thus (ii) for motivating the theoretical significance of their distinctness. By closing the formal gap between the apriori and aposteriori, while demonstrating the import of their distinctness to the conditions on intentional content, the reasons adducing in favor of the robust significance of the distinction may

thus be found to stand on firmer grounds.

Part IV: Epistemic Modality and the Philosophy of Mathematics

Chapter 9

Epistemic Modality, Necessitism, and Abstractionism

Modal notions have been availed of, in order to argue in favor of nominalist approaches to mathematical ontology. Field (1989) argues, for example, that mathematical modality can be treated as a logical consistency operator on a set of formulas comprising an empirical theory, such as Newtonian mechanics, in which the mathematical vocabulary has been translated into the vocabulary of physical geometry.[1] Putnam (1967), Parsons (1983), Chihara (1990), and Hellman (1993) argue that intensional models both of first- and second-order arithmetic and of set theory motivate an eliminativist approach to mathematical ontology. On this approach, reference to mathematical objects can be eschewed, and possibly the mathematical structures at issue are nothing.[2]

This essay aims to provide modal foundations for mathematical platonism, i.e., the proposal that mathematical terms for sets; functions; and the natural, rational, real, and imaginary numbers refer to abstract – necessarily non-concrete – objects. Intensional constructions of arithmetic and set theory have been developed by, inter alia, Fine (1981); Parsons (op. cit.); Shapiro (1985); Myhill (1985); Reinhardt (1988); Nolan (2002); Lin-

---

[1] For a generalization of Field's nominalist translation scheme to the differential equations in the theory of General Relativity, see Arntzenius and Dorr (2012).

[2] For further discussion of modal approaches to nominalism, see Burgess and Rosen (1997: II, B-C) and Leng (2007; 2010: 258).

nebo (2013); and Studd (2013). Williamson (2013; 2014) emphasizes that mathematical languages are extensional, although in Williamson (2016) he argues that Orey sentences, such as the generalized continuum hypothesis – $2^{\aleph_\alpha} = \aleph_{\alpha+1}$ – which are currently undecidable relative to the axioms of the language of Zermelo-Fraenkel Set Theory with choice as augmented by large cardinal axioms, are yet possibly decidable.[3] Chapters **9** and **11** argue that the epistemic interpretation of two-dimensional intensional semantics provides a novel approach to the epistemology of mathematics, such that if the decidability of mathematical axioms is epistemically possible, then their decidability is metaphysically possible.[4] Epistemic mathematical modality, suitably constrained, can thus serve as a guide to metaphysical mathematical modality.[5] Hamkins and Löwe (2007; 2013) argue that the modal logic of set-forcing extensions and the corresponding logic for their ground models satisfy at least S4.2, i.e., axioms K $[\Box(\phi \to \psi) \to (\Box\phi \to \Box\psi)]$; T $(\Box\phi \to \phi)$; 4 $(\Box\phi \to \Box\Box\phi)$; and G $(\Diamond\Box\phi \to \Box\Diamond\phi)$. While the foregoing approaches are consistent with realism about mathematical objects, they are nevertheless not direct arguments thereof. The aim of this essay is to redress the foregoing lacuna, and thus to avail of the resources of modal ontology and epistemology in order to argue for the reality of mathematical entities and truth.

In Section **2**, I outline the elements of the abstractionist foundations of mathematics. In Section **3**, I examine Hale and Wright (2009)'s objections to the merits and need, in the defense of mathematical platonism and its epistemology, of the thesis of Necessitism, underlying the thought that whatever can exist actually does so. The Necessitist thesis is codified by the Barcan formula (cf. Barcan, 1946; 1947), and states that possibly if there is something which satisfies a condition, then there is something such that it possibly satisfies that condition: $\Diamond\exists x\phi x \to \exists x\Diamond\phi x$. I argue that Hale and Wright's

---

[3]Compare Reinhardt (1974) on the imaginative exercises taking the form of counterfactuals concerning the truth of undecidable formulas. See Maddy (1988,b), for critical discussion.

[4]The epistemic interpretation of two-dimensional intensional semantics is first advanced in Chalmers (1996; 2004).

[5]See Section **4**, for further discussion. Gödel (1951: 11-12) anticipates a similar distinction between epistemic and metaphysical readings of the determinacy of mathematical truths, by distinguishing between mathematics in its subjective and objective senses. The former concerns decidable formulas, while the latter records the values of formulas defined, owing to the incompleteness theorems, in a variant of the language augmented by stronger axioms of infinity.

objections to Necessitism as a requirement on admissible abstraction can be answered; and I examine both the role of the higher-order Necessitist proposal in their endorsement of an abundant conception of properties, as well as cardinality issues that arise depending on whether Necessitism is accepted at first- and higher-order. In Section **4**, I provide an account of the role of epistemic and metaphysical modality in explaining the prima facie justification to believe the truth of admissible abstraction principles, and demonstrate how it converges with both Hale and Wright's (op. cit.) and Wright's (2012; 2014) preferred theory of default entitlement rationally to trust the truth of admissible abstraction. Section **5** provides concluding remarks.

9.1 The Abstractionist Foundations of Mathematics

The abstractionist foundations of mathematics are inspired by Frege's (1884/1980; 1893/2013) proposal that cardinal numbers can be explained by specifying an equivalence relation, expressible in the signature of second-order logic and identity, on lower-order representatives for higher-order entities. Thus, e.g., in Frege (1884/1980: 64), the direction of the line, a, is identical to the direction of the line, b, if and only if lines a and b are parallel. In Frege (op. cit.: 68) and Wright (1983: 104-105), the cardinal number of the concept, **A**, is identical to the cardinal number of the concept, **B**, if and only if there is a one-to-one correspondence between **A** and **B**, i.e., there is an injective and surjective (bijective) mapping, R, from **A** to **B**. With Nx: a numerical term-forming operator,

- $\forall\mathbf{A}\forall\mathbf{B}\exists R[[\text{Nx: }\mathbf{A} = \text{Nx: }\mathbf{B} \equiv \exists R[\forall x[\mathbf{A}x \rightarrow \exists y(\mathbf{B}y \wedge Rxy \wedge \forall z(\mathbf{B}z \wedge Rxz \rightarrow y = z))] \wedge \forall y[\mathbf{B}y \rightarrow \exists x(\mathbf{A}x \wedge Rxy \wedge \forall z(\mathbf{A}z \wedge Rzy \rightarrow x = z))]]$.

The foregoing is referred to as 'Hume's Principle'.[6] Frege's Theorem states that the Dedekind-Peano axioms for the language of arithmetic can be de-

---

[6]Frege (1884/1980: 68) writes: 'the Number which belongs to the concept F is the extension of the concept '[equinumerous] to the concept F' (cf. op. cit.: 72-73). Boolos (1987/1998: 186) coins the name, 'Hume's Principle', for Frege's abstraction principle for cardinals, because Frege (op. cit.: 63) attributes equinumerosity as a condition on the concept of number to Hume (1739-1740/2007: Book 1, Part 3, Sec. 1, SB71), who writes: 'When two numbers are so combin'd, as that the one has always an unite answering to every unite of the other, we pronounce them equal . . . '. Frege notes that identity of number

rived from the Hume's Principle, as augmented to the signature of second-order logic and identity.[7] Abstraction principles have further been specified both for the real numbers (cf. Hale, 2000a; Shapiro, 2000; and Wright, 2000), and for sets (cf. Wright, 1997; Shapiro and Weir, 1999; Hale, 2000b; and Walsh, 2016).

The philosophical significance of the abstractionist program consists primarily in its provision of a neo-logicist foundation for classical mathematics, and in its further providing a setting in which to examine constraints on the identity conditions constitutive of mathematical concept possession.[8] The philosophical significance of the abstractionist program consists, furthermore, in its circumvention of Benacerraf's (1973) challenge to the effect that our knowledge of mathematical truths is in potential jeopardy, because of the absence of naturalistic, in particular causal, conditions thereon. Abstraction principles provide an epistemic conduit into our knowledge of abstracta, because a grasp of the right-hand side of the principles – i.e., the equivalence relation on lower-order entities, such as the equinumerosity of two, distinct objects – is explanatory of the left-hand side of the principles – i.e., a grasp of the identity of the numbers of the higher-order entities (the concepts) under which those objects fall. Both Wright (1983: 13-15) and Hale (1987: 10-15) argue, then, that the abstraction principles are epistemically tractable, only if (i) the surface syntax of the principles – e.g., the term-forming operators referring to objects – are a perspicuous guide to their logical form; and (ii) the principles satisfy Frege's (1884/1980: X) context principle, such that the truth of the principles is secured prior to the reference of the terms figuring therein.

---

via bijections is anticipated by the mathematicians, Ernst Schröder and Ernst Kossak, as well Cantor (1883/1996: Sec. 1), who writes: '[E]very well-defined set has a determinate power; two sets have the same power if they can be, element for element, correlated with one another reciprocally and one-to-one', where the power [Anzahl] of a set corresponds to its cardinality (cf. Cantor, 1895/2007: 481).

[7]Cf. Dedekind (1888/1996) and Peano (1889/1967). See Wright (1983: 154-169) for a proof sketch of Frege's theorem; Boolos (1987) for the formal proof thereof; and Parsons (1964) for an incipient conjecture of the theorem's validity.

[8]Shapiro and Linnebo (2015) prove that Heyting arithmetic can be recovered from Frege's Theorem. Criteria for consistent abstraction principles are examined in, inter alia, Hodes (1984,a); Hazen (1985); Boolos (1990/1998); Heck (1992); Fine (2002); Weir (2003); Cook and Ebert (2005); Linnebo and Uzquiano (2009); Linnebo (2010); and Walsh (op. cit.).

9.2 Abstraction and Necessitism

9.2.1 Hale and Wright's Arguments against Necessitism

One crucial objection to the abstractionist program is that – while abstraction principles might provide a necessary and sufficient truth-condition for our grasp of the concepts of mathematical objects – an explanation of the actual truth of the principles has yet to be advanced (cf. Eklund, 2006; 2016). In response, Hale and Wright (2009: 197-198) proffer a tentative endorsement of an 'abundant' conception of properties, according to which fixing the sense of a predicate will be sufficient for predicate reference. Eklund (2006: 102) suggests, by contrast, that one way for the truth of the abstraction principles to be explained is by presupposing what he refers to as a 'Maximalist' position concerning the target ontology.[9] According to the ontological Maximalist position, if it is possible that a term has a certain extension, then actually the term does have the designated extension.

Hale and Wright (op. cit.) raise two issues for the ontological Maximalist proposal. The first is that ontological Maximalism is committed to a proposal that they take to be independently objectionable, namely ontological Necessitism (185). Hale and Wright (op. cit.) raise a similar contention to the effect that actual, and not *merely possible*, reference is what the abstractionist program intends to target; and that Maximalism and Necessitism, so construed, are purportedly silent on the status of ascertaining when the possibilities at issue are actual.

The second issue that Hale and Wright find with Maximalism is that it misconstrues the demands that the abstractionist program is required to address. The abstractionist program is supposed to be committed to ontological Maximalism, because the possibility that a term has a certain extension will otherwise not be sufficient for the success of the term's reference. It is further thought that, without an appeal to Maximalism, and despite the actuality of successful mathematical predicate reference, there are yet possible situations in which the mathematical predicates still do not refer (193). In response, they note that no 'collateral metaphysical assistance' – such as ontological Maximalism would be intended to provide – is necessary in order to explain the truth of abstraction principles (op. cit.). Rather, there is prima facie, default entitlement rationally to trust that the abstraction principles are ac-

---

[9]For further discussion of ontological Maximalism, see Hawley (2007) and Sider (2007: IV).

tually true, and such entitlement is sufficient to foreclose upon the risk that possibly the mathematical terms therein do not refer (192).

In the remainder of this section, I will argue that Hale and Wright's objections to Necessitism and the ontological Maximalist approach to admissible abstraction both can be answered, and in any case are implicit in their endorsement of the abundant conception of properties. In the following section, I address their second contention, and I argue for the fundamental role that Maximalism and Necessitism can play in warranting the truth of candidate abstractions and mathematical platonism.

The principle of the necessary necessity of being (NNE) can be derived from the Barcan formula.[10] NNE states that necessarily everything is necessarily such that there is something to which it is identical; $\Box \forall x \Box \exists y (x = y)$. Informally, necessarily everything has necessary being, i.e. everything is something, even if contingently non-concrete. Williamson (2013: 6.1-6.4) targets issues for haecceity comprehension, if the negations of the Barcan formula and NNE are true at first-order, and thus for objects. With regard to properties and relations at higher-order, Williamson's arguments have targeted closure conditions, given a modalized interpretation of comprehension principles (op. cit.). The latter take the form, $\exists X \Box \forall x (Xx \iff \phi)$, with x an individual variable which may occur free in ϕ and X a monadic first-order predicate variable which does not occur free in ϕ (262).[11] He targets, in particular, the principle of mathematical induction – with s a successor function and the quantifier ranging over the natural numbers: $\forall X[[X0 \wedge \forall n(Xn \rightarrow Xsn)] \rightarrow \forall n(Xn)]$ – and notes that instances of mathematical induction – e.g., for a an individual constant, $\exists R[[Ra0 \wedge \forall n(Ran \rightarrow Rasn)] \rightarrow \forall n(Ran)]$ – presuppose, for their derivation, the validity of instances of the higher-order modal comprehension scheme: e.g., $\exists X \Box \forall n(Xn \iff Ran)$ (283-284). The foregoing provides prima facie abductive support for the requirement of Necessitism in the practice of mathematics. The constitutive role of the Necessitist modal comprehension scheme in the principle of mathematical induction answers Hale and Wright's first contention against the Necessitist commitments of ontological Maximalism.

Williamson refers to the assignments for models in the metaphysical setting as universal interpretations (59). The analogue for logical truth occurs

---

[10]Cf. Williamson (2013: 38).

[11]The contingentist, by contrast, can – by rejecting the Barcan formula – countenance only 'intra-world' comprehension principles in which the modal operators and iterations thereof take scope over the entire formula; e.g. $\Diamond \forall x \exists X(Xx \iff \phi)$ (cf. Sider, 2016: 686).

when a truth is metaphysically universal, i.e., if and only if its second-order universal generalization is true on the intended interpretation of the metalanguage (200). The connection between truth-in-a-model and truth simpliciter is then that – as Williamson puts it laconically – when 'the framework at least delivers a condition for a modal sentence to be true in a universal interpretation, we can derive the condition for it to be true in the intended universal interpretation, which is the condition for it to be true simpliciter' (op. cit.).

One of the crucial interests of the metaphysical universality of formulas is that the models in the class need not be pointed, in order to countenance the actuality of the possible formulas defined therein.[12] Rather, the class of true, possible propositions generated by the metaphysically universal possible formulas is sufficient for the formulas actually to be true (268-269). To see this, let the quantifier for worlds in modal formulas which are true simpliciter range over real metaphysically possible worlds. Suppose that the Barcan formula is true such that, possibly if there is something which satisfies a condition then there is something such that possibly it satisfies that condition. Then, because there is something which possibly does so, Williamson writes that 'since whatever is is, whatever is actually is: if there is something, then there actually is such a thing' (23). When the quantifier ranges, then, over possible worlds treated as objects, the worlds are thus actual. So – by contrast to modal formulas which are true-in-a-model – the objects, i.e. actual worlds, which fall within the range of the quantifiers in the metaphysically universal formulas which are true simpliciter can explain why such formulas are actual.

The constitutive role of metaphysical universality in bridging the necessary necessity of being with the actuality thereof answers Hale and Wright's contention that the interaction between the possible and actual truth of abstraction principles has yet to be accounted for.[13]

---

[12]That the models are unpointed is noted in Williamson (2013: 100).

[13]Cook (2016: 398) demonstrates how formally to define modal operators within Hume's Principle, i.e. the consistent abstraction principle for cardinal numbers. Necessitist Hume's Principle takes the form: $\Box\forall X,Y[\#(X) =_\Box \#(Y) \iff X \approx Y]$, where X and Y are second-order variables, $\#$ is a numerical term-forming operator, $\approx$ is a bijection, and for variables, x,y, of arbitrary type 'x $=_\Box$ y $\iff \exists z[z = x \land z = y \land \Box\exists w(w = z)]$'. See Cook (op. cit.) for further discussion.

9.2.2 Hale on the Necessary Being of Purely General Properties and Objects

Note, further, that the abundant conception of properties endorsed by Hale and Wright depends upon the Necessitist Thesis, and the truth of ontological Maximalism thereby. Hale writes: '[I]t is sufficient for the *actual* existence of a property or relation that there *could* be a predicate with appropriate satisfaction conditions ... *purely general* properties and relations exist as a matter of (absolute) *necessity*', where a property is purely general if and only if there is a predicate for which, and it embeds no singular terms (Hale, 2013b: 133, 135; see also 2013a: 99-100).[14]

Hale argues for the necessary necessity of being for properties and propositions as follows (op. cit.: 135; 2013b: 167). Suppose that p refers to the proposition that a property exists, and that q refers to the proposition that a predicate for the property exists. Let the necessity operator be defined as a counterfactual with an unrestricted, universally quantified antecedent, such that, for all propositions, ψ: $[\Box\psi \iff \forall\phi(\phi \Box\rightarrow \psi)]$ (135).[15] On the abundant conception of properties, $\Box[p \iff \diamond q]$. Intuitively: Necessarily, there is a property if and only if possibly there is a predicate for that property. Given the counterfactual analysis of the modal operator: For all propositions about a property, if there were a proposition specifying a predicate s.t. the property is in the predicate's extension, then there would be that property. Conversely, for all propositions specifying a predicate s.t. a property is in the predicate's extension, if there were that property, then there would be a predicate which refers to that property.

From '$\Box[p \iff \diamond q]$', one can derive both '$p \iff \diamond q$', and – by the rule, RK – the necessitation thereof, '$\Box p \iff \Box\diamond q$' (op. cit.). By the B axiom in S5, $\diamond q \iff \Box\diamond q$ (op. cit.). So, '$\Box\diamond q \iff \diamond q$'; '$\diamond q \iff p$'; and '$\Box\diamond q \iff \Box p$'. Thus – by transitivity – '$p \iff \Box p$' (op. cit.); i.e., all propositions about properties are necessarily true, such that the corresponding properties have necessary being. By the 4 axiom in S5, $\Box p \iff \Box\Box p$; so, the necessary being of properties and propositions is itself

---

[14]Cook (op. cit.: 388) notes the requirement of Necessitism in the abundant conception of properties, although does not discuss points at which Williamson's and Hale's Necessitist proposals might be inconsistent. The points of divergence between the two variations on the proposal are examined below.

[15]Proponents of the translation from modal operators into counterfactual form include Stalnaker (1968/1975), McFetridge (1990: 138), and Williamson (2007).

necessary. Given the endorsement of the abundant conception of properties – Hale and Wright are thus committed to higher-order necessitism, i.e., the necessary necessity of being.

Hale (2013b) endeavors to block the ontological commitments of the Barcan formula and its converse by endorsing a negative free logic. Thus, in the derivation:

Assumption,

1. $\Box\forall x[F(x)]$.

By $\Box$-elimination,

2. $\forall x[F(x)]$.

By $\forall$-elimination,

3. $F(x)$.

By $\Box$-introduction,

4. $\Box[F(x)]$.

By $\forall$-introduction,

5. $\forall\Box[F(x)]$.

By $\rightarrow$-introduction,

6. $\Box\forall x[F(x) \rightarrow \forall\Box[F(x)]$,

Hale imposes an existence-entailing assumption in the inference from lines (2) to (3), i.e.

'(Free$\forall$-Elimination) From $\forall x[A(x)]$, together with an existence-entailing premise $F(t)$, we may infer $A(t)$ where t can be any term' (op. cit.: 208-209).

Because the concept of, e.g., cardinal number is defined by abstraction principles which are purely general because they embed no singular terms, the properties – e.g., the concepts – of numbers are argued to have necessary being. The necessary being of the essential properties of number – i.e., higher-order Necessitism about purely general properties – is argued then to explain in virtue of what abstract objects such as numbers and functions have themselves necessary being (176-177). Thus the necessary being of predicate sense for the concept of number can both suffice for and explain the necessary being of predicate reference, i.e. the necessary existence of numbers.

By contrast, essential properties defined by theoretical identity statements, which if true are necessarily so, do embed singular terms and are thus not purely general. So, the essential nature of water, i.e., the property 'being comprised of one oxygen and two hydrogen molecules', has contingent being, explaining in virtue of what samples of water have contingent being (216-217).

Objections

One argument against the above approach is that there is an asymmetry in the order of explanation with regard to whether properties or objects are the source of the contingency with which the lower- and higher-order entities have being. Concrete individuals susceptible to being defined via theoretical identity statements have, as noted, contingent being because the identity statements characterizing the essential properties – e.g., the concepts – of those individuals are not purely general, because they embed singular terms (op. cit.). However, for concrete individuals which cannot be defined via theoretical identity statements, the order of explanation is reversed. The contingency of the essential properties of concrete individuals, such as stars, depends upon the existence-entailing premise noted above (217). Thus, because concrete first-order objects such as stars can be rendered non-concrete, they have contingent being. Higher-order Contigentism is argued, then, to take into effect, because the essential properties thereof exist only when their corresponding objects do so. An explanation of the asymmetry in the foregoing order of explanation of first-and higher-order Contingentism is, however, wanting.

The necessary being of purely general properties – e.g., the possible predicate sense for the concept of number – is argued to hold with absolute modality, as the consequent of a counterfactual with an unrestictedly general antecedent. As Hale notes, absolute modalities are therefore logical modalities (100). Thus, a crucial virtue of the logical necessity with which the concept of number – and thus the numbers themselves – have being, is that it provides a further vindication of the Neo-logicist thesis, that second-order logic and identity as augmented by the purely general abstraction principles is sufficient to derive the axioms of mathematical languages such as Peano arithmetic. The Neo-logicist thesis is qualified, however, by the caveat that abstraction principles are non-logical truths, i.e., they are not true on all interpretations of the values of the variables comprising their syntactic form. By arguing that purely general properties have necessary being and defining the necessity of being of purely general properties as logical necessity, abstraction principles – while not themselves logically true – can thus still hold of logical necessity. The logical necessity of abstraction principles – owing to the purely general predicates comprising their form – can thus serve to explain another sense in which the abstractionist foundations provide a logical reduction of mathematical truths. Mathematical truths will, on the above

approach, be derivable from second-order logic with identity and implicit definitions which are true of logical necessity.

A second objection to the foregoing is that Hale takes the modal status of the being with which both logical and non-logical properties exist to be equivalent. Following Fine (2005), he notes that there is a distinction between 'unworldly' or 'transcendental' truths about individuals, which are true, in Fine's phrasing, 'on the basis of [their] logical form alone and without regard to the circumstances' (Fine, op. cit.: 324). An example of a transcendental truth is that 'Hypatia is self-identical', i.e., '$\Box \exists x(x = H \wedge \Box x = x)$'. A 'worldly' or 'necessary' truth is, by contrast, one whose truth-value is defined in a world; e.g., that 'water = H20' or – if one were to augment one's language with an existence predicate beyond the quantifiers – that 'Socrates exists or does not exist' (op. cit.).

However, Hale draws, as noted, no similar distinction between the modal status of a sentence true in virtue of its logical form, and a worldly sentence whose truth depends on non-logical values of its constituent variables, e.g., a truth of physics (Hale, op. cit.: 215). Hale (2000/2001: 415) notes Wright's (2001: 315) argument that the conjunction of two predicates, e.g., being blue and being self-identical, is equivalent to one of the conjuncts, e.g., being blue. Thus, the predicate for the property, being self-identical, cannot be purely general. He argues, thus, that the truth of 'this star is self-identical' depends on the concrete existence of the star, such that the logical property, being self-identical, has contingent being. Thus, the status of the being of logical properties is contingent, in the same manner that the essential property, being, e.g., H20, depends on concrete instances of water.

A problematic consequence of the foregoing is that the logical necessity of a formula will thus depend on whether the predicates therein are purely general by embedding no singular terms, rather than on whether the formula at issue is a logical truth. It might be replied that Hale is following Frege in defining one of the constitutive marks of logical truths as consisting in their generality – e.g., the generality of their application (1893/2013: XV; 1897/1997), as well as whether the formula is a true universal generalization (op. cit.: §8-9) – rather than Tarski's (1936/1983: 415-417) definition of a logical truth as a formula true in virtue of its logical form and thus whose truth is invariant under permutation of the values of the variables which replace the non-logical constants therein. However – even if not a purely general property because it embeds singular terms – the reflexivity of identity is a logical law, because – as Frege himself writes of reflexivity – 'the value

of this function is always the True, whatever we take as argument' (Frege, 1891/1997: 23).[16]

A final objection concerns the necessary being of different types of numbers. While an abstraction principle for cardinal numbers can be specified using only purely general predicates – i.e., Hume's Principle –abstraction principles for imaginary and complex numbers have yet to be specified. Shapiro (2000) provides an abstraction principle for the concepts of the reals by simulating Dedekind cuts, where abstraction principles are provided for the concepts of the cardinals, natural numbers, integers, and rational numbers, from which the reals are thence defined: Letting F,G, and R denote rational numbers, $\forall F,G[\mathbf{C}(F) = \mathbf{C}(G) \iff \forall R(F \leq R \iff G \leq R)]$. [17] Hale's (2000/2001) own definition of the concept of the reals is provided relative to a domain of quantities. The quantities are themselves taken to be abstract, rather than physical, entities (409). The quantitative domain can thus be comprised of both rational numbers as well as the abstracts for lengths, masses, and points.[18] The reals are then argued not to be numbers, but rather quantities defined via an abstraction principle which states that a set of rational numbers in one quantitative domain is identical to a set of rational numbers in a second quantitative domain if and only if the two domains are isomorphic (407).[19] Hale argues, then, that it is innocuous for the real abstraction principle to be conditional on the existence of at least one quantitative domain, because the rational numbers can be defined, similarly as on Shapiro's approach, via cut-abstractions and abstractions on the integers, naturals, and cardinals. Thus, the reals can be treated as abstracts derived from purely general abstraction principles, and are thus possessed of necessary being. However, abstraction principles for imaginary numbers such as $i = \sqrt{-1}$, and complex numbers which are defined as the sum of a

---

[16]For further discussion of Frege's treatment of logic as a 'purely general science', see Dummett (1991: 224-225); MacFarlane (2002); Burge (2005: 133, 137-138); and Blanchette (2012: 1.1, 1.22, 3.6). For further discussion of Tarski's account of logical truth, see Etchemendy (1990), McGee (1992, 1996), Gómez-Torrente (1996), Chihara (1998), Feferman (1999), and Sher (2008).

[17]See Dedekind (1872/1996: Sec. 4), for the cut method for the definition of the reals.

[18]An abstraction principle for lengths, based on the equivalence property of congruence relations on intervals of a line, or regions of a space, is defined in Shapiro and Hellman (2015: 5, 9). Shapiro and Hellman provide, further, an abstraction principle for points, defined as comprising, respectively, the left- and right-ends of intervals (op. cit.: 5, 10-12).

[19]Cf. Hale (op. cit.: 406-407), for the further conditions that the domains are required to satisfy.

real number and a second real multiplied by i, have yet to be accounted for. The provision of an abstraction principle for complex numbers would, in any case, leave open the inquiry into how, e.g., complex-valued wave functions might interact with physical ontology; e.g., whether such functions might be metaphysically fundamental entities which serve to represent physical fields in higher-dimensional spacetime, and whether or how the domain of the functions, i.e., a real-valued configuration space for particles, might relate to the higher-dimensional, complex-valued wave function (cf. Simons, 2016; Ney, 2013; Maudlin, 2013).

The modality in the Barcan-induced Necessitist proposal at first- and higher-order is, as noted, interpreted metaphysically rather than logically, and thus incurs no similar issues with regard to the interaction between purely general properties, logical properties, and concrete entities. Further, because true on its second-order universal generalization on its intended, metaphysical interpretation, the possible truth-in-a-model of the relevant class of formulas is, as discussed in Section **3.1**, thus sufficient for entraining the actual truth of the relevant formulas.

9.2.3   Cardinality and Intensionality

An interesting residual question concerns the status of the worlds, upon the translation of modal first-order logic into the non-modal first-order language.[20] Fritz (op. cit.) notes that a world can be represented by a predicate, in the latter.[21] However, whether objects satisfy the predicate can vary from point to point, in the non-modal first-order class of points.[22] Another issue is that modal propositional logic is equivalent only to the bisimulation-invariant fragments of both first-order logic and fixed-point monadic second-order logic, rather than to the full variants of either logic (cf. van Benthem, 1983; Janin and Walukiewicz, 1996). Thus, there cannot be a faithful translation from each modal operator in modal propositional logic into a predicate of full first- or monadic higher-order logic.

One way to mitigate the foregoing issues might be by arguing that the

---

[20]Thanks here to xx, for discussion.

[21]For further discussion of the standard translation between propositional modal and first-order non-modal logics, see Blackburn et al. (2001: 84).

[22]Suppose that the model is defined over the language of second-order arithmetic, such that the points in the model are the ordinals. A uniquely designated point might then be a cardinal number whose height is accordingly indexed by the ordinals.

language satisfies real-world rather than general validity, such that necessarily the predicate will be satisfied only at a designated point in a model – intuitively, the analogue of the concrete rather than some merely possible world, simulating thereby the translation from possibilist to actualist discourse (cf. Fine, op. cit.: 211,135-136, 139-140, 154, 166-168, 170-171) – by contrast to holding of necessity as interpreted as satisfaction at every point in the model.[23] The reply would be consistent with what Williamson refers to as 'chunky-style necessitism' which validates the following theorems: where the predicate $C(x)$ denotes the property of being grounded in the concrete and $P(x)$ is an arbitrary predicate, (a) '$\forall x \Diamond C(x)$', yet (b) '$\Box \forall x [P(x_1, \ldots, x_n) \to (Cx_1, \ldots, Cx_1)]$' (325-332). Williamson (33, fn.5) argues, however, in favor of general, rather than real-world validity. A second issue for the reply is that principle (b), in the foregoing, is inconsistent with Williamson's protracted defense of the 'being constraint', according to which $\Box \forall x \Box [P(x_1, \ldots, x_n) \to \exists y (x = y)]$, i.e. if x satisfies a predicate, then x is something, even if possibly non-concrete (148).

A related issue concerns the translation of modalized, variable-binding, generalized quantifiers of the form:

'there are n objects such that . . . ',

'there are countably infinite objects such that . . . ',

'there are uncountably infinite objects such that . . . ' (Fritz and Goodman, 2017).

The generalized quantifiers at issue are modalized and consistent with first-order Necessitism, because the quantifier domains include all possible – including contingently non-concrete – objects. It might be argued that the translation is not of immediate pertinence to the ontology of mathematics, because the foregoing first-order quantifiers can be restricted such that they range over only uncountably infinite *necessarily* non-concrete objects – i.e. abstracta – by contrast to ranging unrestrictedly over all modal objects, including the contingently non-concrete entities induced via the Barcan formula – i.e., the 'mere possibilia' that are non-concrete as a matter of contingency; e.g., the possible star and the possible nubula which are both actual objects and yet are such that a carpenter actually cannot, as a matter of pairwise incompossibility, render both concrete. However, the Necessitist thesis can be valid even in the quantifier domain of a first-order language restricted to necessarily non-concrete entities. If, e.g., a mathematician takes, despite

---

[23]See ftn. 15.

141

iterated applications of set-forming operations, the cumulative hierarchy of sets to have a fixed cardinal height, then the first-order Necessitist thesis will still be valid, because all possible objects will actually be still something.

The first-order Necessitist proposal engendered by taking the height of the cumulative hierarchy to be fixed is further consistent with the addition to the first-order language of additional intensional operators – such as those introduced by Hodes (1984b) – in order to characterize the indefinite extensibility of the concept of set; i.e., that despite unrestricted universal quantification over all of the entities in a domain, another entity can be defined with reference to, and yet beyond the scope of, that totality, over which the quantifier would have further to range.[24] First-order Necessitism is further consistent with the relatively expanding domains induced by Bernays' (1942) Theorem. Bernays' Theorem states that class-valued functions from classes to sub-classes are not onto, where classes are non-sets (cf. Uzquiano, 2015a: 186-187). So, the cardinality of a class will always be less than the cardinality of its sub-classes. Suppose that that there is a generalization of Bernays' theorem, such that the non-sets are interpreted as possible objects. Thus, the cardinality of the class of possible objects will always be less than the cardinality of the sub-classes in the image of its mapping. Given iterated applications of Bernays' theorem, the cardinality of a domain of non-sets is purported then not to have a fixed height.

In both cases, however, the addition of Hodes' intensional operators permits there to be multiple-indexing in the array of parameters relative to which a cardinal can be defined, while the underlying logic for metaphysical modality can be S5, partitioning the space of worlds into equivalence classes. So, both the intensional characterization of indefinite extensibility and the generalization of Bernays' Theorem to possible objects are consistent with the first-order Necessitist proposal that all possible objects are actual, and so the cardinality of the target universe is fixed.[25]

---

[24]The concept of indefinite extensibility is introduced by Dummett (1963/1978), in the setting of a discussion of the philosophical significance of Gödel's (1931) first incompleteness theorem. See the essays in Rayo and Uzquiano (2006); Studd (op. cit.); Dever (ms); and Author (ms), for further discussion.

[25]Note that the proposal that the cardinality of the cumulative hierarchy of sets is fixed, despite continued iterated applications of set-forming operations, is anticipated by Cantor (1883/1996: Endnote [1]). Cantor writes: 'I have no doubt that, as we pursue this path ever further, we shall never reach a boundary that cannot be crossed, but that we shall also never achieve even an approximate conception of the absolute [. . .] The absolutely infinite sequence of numbers thus seems to me to be an appropriate symbol of the absolute;

Fritz and Goodman suggest that a necessary condition on the equivalence of propositions is that they define the same class of models (op. cit.: 1.4). The proposed translation of the modalized generalized quantifiers would be Contingentist, by taking (NNE) to be invalid, such that the domain in the translated model would be comprised of only possible concrete objects, rather than the non-concrete objects as well (op. cit.).

Because of the existence of non-standard models, the generalized quantifier that 'there are countably infinitely many possible ...' cannot be defined in first-order logic. Fritz and Goodman note that generalized quantifiers ranging over countably infinite objects can yet be simulated by enriching one's first-order language with countably infinite conjunctions. On the latter approach, finitary existential and universal quantifiers can be defined as the countably infinite conjunction of formulas stating that, for all natural numbers n, 'there are n possible ...' (2.3).

Crucially, however, there are some modalized generalized quantifiers that cannot be similarly paraphrased – e.g., 'there are uncountably infinite possible objects s.t. ...' – and there are some modalized generalized quantifiers that cannot even be defined in first-order languages – e.g. '*most* possible objects s.t. ...' (2.4-2.5)

In non-modal first-order logic, it is possible to define generalized quantifiers which range over an uncountably infinite domain of objects, by augmenting finitary existential and universal quantifiers with an uncountably infinite stock of variables and an uncountably infinite stock of conjunctions of formulas (2.4).[26] Fritz and Goodman note, however, that the foregoing

in contrast the infinity of the first number-class (I) [i.e., the first uncountable cardinal, $\aleph_0$ – HK], which has hitherto sufficed, because I can I consider it to be a graspable idea (not a representation), seems to me to dwindle into nothingness by comparison' (op. cit.; cf. Cantor, 1899/1967).

[26]Uncountable cardinals can be defined as follows. For cardinals, x,a,C, let C⊆a be closed unbounded in a, if it is closed [if x < C and $\bigcup$(C∩a) = a, then a∈C] and unbounded ($\bigcup$C = a) (Kanamori, 2012: 360). A cardinal, S, is stationary in a, if, for any closed unbounded C⊆a, C∩S ≠ ∅ (op. cit.). An ideal is a subset of a set closed under countable unions, whereas filters are subsets closed under countable intersections. A cardinal κ is regular if the cofinality of κ – comprised of the unions of sets with cardinality less than κ – is identical to κ. For models A,B, and conditions ϕ, an elementary embedding, j: A → B, is such that $\phi\langle a_1, \ldots, a_n \rangle$ in A if and only if $\phi\langle j(a_1), \ldots, j(a_n) \rangle$ in B (363). A measurable cardinal is defined as the ordinal denoted by the critical point of j, crit(j) (Koellner and Woodin, 2010: 7). Measurable cardinals are inaccessible (Kanamori, op. cit.). Uncountable regular limit cardinals are weakly inaccessible (op. cit.). A strongly inaccessible cardinal is regular and has a strong limit, such that if $\lambda < \kappa$, then $2^\lambda < \kappa$

would require that the quantifiers bind the uncountable variables 'at once', s.t. they must have the same scope. The issue with the proposal is that, in the setting of modalized existential quantification over an uncountably infinite domain, the Contingentist paraphrase requires that bound variables take different scopes, in order to countenance the different possible sets that can be defined in virtue of the indefinite extensibility of cardinal number (op. cit.).

In order to induce the Contingentist paraphrase, Fritz and Goodman suggest defining 'strings of infinitely many existential and universal quantifiers', such that a modalized, i.e. Necessitist, generalized quantifier of the form, 'there are uncountably infinite possible ...' can be redefined by an uncountably infinite sequence of finitary quantifiers with infinite variables and conjunction symbols of the form:

'Possibly for some x_1, possibly for some x_2, *etc.*: x_1,x_2,*etc.* are pairwise distinct and are each possibly ...',

where *etc.* denotes an uncountable sequence of, respectively, 'an uncountable string of interwoven possibility operators and existential quantifiers', and an 'uncountable string of variables' (op. cit.).

An argument against the proposed translation of the quantifier for there being uncountably infinite possible objects is that it is contentious whether an uncountable sequence of operators or quantifiers has a definite meaning [cf. Williamson (2013: 7.7)]. Thus, e.g., while negation can have a determinate truth condition which specifies its meaning, a string of uncountably infinite negation operators will similarly have determinate truth conditions and yet not have an intuitive, definite meaning (357). One can also define a positive or negative integer, x, such that *s*x is interpreted as the successor function, x+1, and *p*x is interpreted as the inverse function, x-1. However, an infinitary expression consisting in uncountable, alternating iterations of the successor and inverse functions – *spsps*... x – will similarly not have a definite meaning (op. cit.). Finally, one can define an operator O_i mapping truth conditions for an arbitrary formula A to the truth condition, *p*, of the formula $\diamond\exists x_i(Cx_i \wedge A)$, with Cx being the predicate for being concrete (258). Let the operators commute s.t. such that O_iO_j iff O_jO_i, and be idempotent such that O_iO_i iff O_i (op. cit.). A total ordering of truth conditions defined by an infinite sequence of the operators can be defined,

---

(op. cit.). For the foregoing and further definitions, see Koellner and Woodin (op.cit.); Kanamori (op. cit.), and Woodin (2009, 2010).

s.t. that the relation is reflexive, anti-symmetric, transitive, and connected [∀x,y(x≤y ∨ y≤x)] (op. cit.). However, total orders need not have a least upper bound; and the sequence, $O_iO_iO_i\ldots(p)$, would thus not have a non-arbitrary, unique value (op. cit.). The foregoing might sufficiently adduce against Fritz and Goodman's Contingentist paraphrase of the uncountable infinitary modalized quantifier.

The philosophical significance of the barrier to a faithful translation from modal first-order to extensional full first-order languages, as well as a faithful translation from modalized, i.e. Necessitist, generalized quantifiers to Contingentist quantification, is arguably that the modal resources availed of in the abstractionist program might then be ineliminable.

9.3 Epistemic Modality, Metaphysical Modality, and Epistemic Utility and Entitlement

In this section, I address, finally, Hale and Wright's second issue with the role of Necessitism in guaranteeing that the possible truth of abstraction principles provides warrant for the belief in their actual truth. As noted, Hale and Wright argue, against the foregoing approach, that there is non-evidential entitlement rationally to trust that acceptable abstraction principles are true, and thus that the terms defined therein actually refer. In response, I will proceed by targeting the explanation in virtue of which there is such epistemic, default entitlement. I will outline two proposals concerning the foregoing grounding claim – advanced, respectively, in Chapter **11** and by Wright (2012; 2014) – and I will argue that the approaches converge.

Wright's elaboration of the notion of rational trust, which is intended to subserve epistemic entitlement, appeals to a notion of 'expected epistemic utility' in the setting of decision theory (2014: 226, 241). In order better to understand this notion of expected epistemic utility, we must be more precise.

There are two, major interpretations of (classical) expected utility.[27] A

---

[27]For an examination of non-classical utility measures, see Buchak (2014). Non-classical utility measures are intended to describe the innocuous rationality with which an agent's expected utility might diminish with the order of the bets she might pursue. In the latter case, her expected utility will then be sensitive to her propensity to take risks relative to

model of decision theory is a tuple $\langle A,O,K,V\rangle$, where A is a set of acts; O is a set of outcomes; K encodes a set of counterfactual conditionals, where an act from A figures in the antecedent of the conditional and O figures in the conditional's consequent; and V is a function assigning a real number to each outcome. The real number is a representation of the value of the outcome. In evidential decision theory, the expected utility of an outcome is calculated as the product of the agent's credence, conditional on her action, by the utility of the outcome. In causal decision theory, the expected utility of an outcome is calculated as the product of the agent's credence, conditional on both her action and background knowledge of the causal efficacy thereof, by the utility of the outcome.

First, because background knowledge concerning the causal efficacy of one's choice of acts is presumably orthogonal to the non-evidential rational trust to believe that mathematical abstraction principles are true, I will assume that the notion of expected epistemic utility theory that Wright (op. cit.) avails of relies only on the subjective credence of the agent, multiplied by the utility that she assigns to the outcome of the proposition in which she's placing her rational trust. Thus expected epistemic utility in the setting of decision-theory will be calculated within the (so-called) evidential, rather than causal, interpretation of the latter.

Second, there are two, major interpretations concerning how to measure the subjective credences of an agent. The philosophical significance of this choice point is that it bears directly on the very notion of the *epistemic utility* that an agent's beliefs will possess. So, e.g., according to pragmatic accounts of the accuracy of one's partial beliefs, one begins by defining a preference ordering on the agent's space of acts and outcomes. If the preference ordering is consistent with the Kolmogorov axioms[28], then one can set up a representation theorem from which the agent's subjective probability and utility measures (i.e., their expected utility measure) can be derived.[29]

---

the total ordering of the gambles, such that she can have a preference for a sure-gain of .5 units of value, rather than prefer a bet with a 50 percent chance of winning either 0 or 1 units of value.

[28]Namely: normality (which states that the probability of a tautology maps to 1); non-negativity (which states that the probabilty operator must take a non-negative value); additivity (which states that for all disjoint probability densities, the probability of their union is equal to the probability of the first density added to the probability of the second); and conditionalization [which states that the probability of ϕ conditional on ψ equals the probability of the intersection of ϕ and ψ, divided by the probability of ψ].

[29]Cf. Ramsey (1926); Savage (1954); and Jeffrey (1965).

The epistemic utility associated with the pragmatic approach is, generally, utility maximization.

By contrast to the pragmatic approach, the epistemic approach to measuring the accuracy of one's beliefs is grounded in the notion of dominance (cf. Joyce, 1998; 2009). According to the epistemic approach, there is an ideal, or vindicated, probability concerning a proposition's obtaining, and if an agent's subjective probability measure does not satisfy the Kolmogorov axioms, then one can prove that it will always be dominated by a distinct measure; i.e. it will always be the case that a distinct subjective probability measure will be closer to the vindicated world than one's own. The epistemic utility associated with the epistemic approach is thus the minimization of inaccuracy (cf. Pettigrew, 2014).[30]

Wright notes that rational trust subserving epistemic entitlement will be pragmatic, and makes the intriguing point that 'pragmatic reasons are not a special genre of reason, to be contrasted with e.g. epistemic, prudential, and moral reasons' (2012: 484). He provides an example according to which one might be impelled to prefer the 'alleviation of Third world suffering' to one's own 'eternal bliss' (op. cit.); and so presumably has the pragmatic approach to expected utility in mind. The intriguing point to note, however, is that epistemic utility is variegated; one's epistemic utility might consist, e.g., in both the reduction of epistemic inaccuracy and in the satisfaction of one's preferences. Wright concludes that there is thus 'no good cause to deny certain kinds of pragmatic reason the title 'epistemic'. This will be the case where, in the slot in the structure of the reasons for an action that is to be filled by the desires of the agent, the relevant desires are focused on epistemic goods and goals' (op. cit.).

Third, and most crucially: The very idea of expected epistemic utility in the setting of decision theory makes implicit appeal to the notion of possible

---

[30]The distinction between the epistemic (also referred to as the alethic) and the pragmatic approaches to epistemic utility is anticipated by Clifford (1877) and James (1896), with Clifford endorsing the epistemic approach, and James the pragmatic. The distance measures comprising the scoring rules for the minimization of inaccuracy are examined in, inter alia, Fitelson (2001); Leitgeb and Pettigrew (2010); and Moss (2011). A generalization of Joyce's argument for probabilism to models of non-classical logic is examined in Paris (2001) and Williams (2012). A dominance-based approach to decision theory is examined in Easwaran (2014), and a dominance-based approach to the notion of coherence – which can accommodate phenomena such as the preface paradox, and is thus weaker than the notion of consistency in an agent's belief set – is examined in Easwaran and Fitelson (2015).

worlds. The full and partial beliefs of an agent will have to be defined on a probability distribution, i.e. a set of epistemically possible worlds. The philosophical significance of this point is that it demonstrates how Hale and Wright's appeal to default, rational entitlement to trust that abstraction principles are true converges with the modal approach to the epistemology of mathematics advanced in Chapter **11**. The latter proceeds by examining undecidable sentences via the epistemic interpretation of two-dimensional intensional semantics. The latter can be understood as recording the thought that the semantic value of a proposition relative to a first parameter (a context) which ranges over epistemically possible worlds, will constrain the semantic value of the proposition relative to a second parameter (an index) which ranges over metaphysically possible worlds. The formal clauses for epistemic and metaphysical mathematical modalities are as follows:

Let C denote a set of epistemically possibilities, such that $[\![\phi]\!]^c \subseteq C$;

(ϕ is a formula encoding a state of information at an epistemically possible world).

-$\texttt{pri}(x) = \lambda c.[\![x]\!]^{c,c}$;

(the two parameters relative to which x – a propositional variable – obtains its value are epistemically possible worlds. The function from possible formulas to values is thus an intension).

-$\texttt{sec}(x) = \lambda w.[\![x]\!]^{w,w}$

(the two parameters relative to which x obtains its value are metaphysically possible worlds).

Then:

- **Epistemic Mathematical Necessity (Apriority)**

  $[\![\blacksquare\phi]\!]^{c,w} = 1 \Longleftrightarrow \forall c'[\![\phi]\!]^{c,c'} = 1$

  (ϕ is true at all points in epistemic modal space).

- **Epistemic Mathematical Possibility**

  $[\![\diamond_c\phi]\!]^c \neq \emptyset \iff [\![Pr\phi]\!]^c \neq \emptyset \wedge > .5$, else $\langle\emptyset, Pr_c(\phi \mid \emptyset)\rangle$.

  (ϕ might be true if and only if its value is not null and it is greater than .5).

Epistemic mathematical modality is constrained by consistency, and the formal techniques of provability and forcing. A mathematical formula is metaphysically impossible, if it can be disproved or induces inconsistency in a model.

- **Convergence**

  $\forall c \exists w [\![\phi]\!]^{c,w} = 1$

  (the value of x is relative to a parameter for the space of epistemically possible worlds. The value of x relative to the first parameter determines the value of x relative to the second parameter for the space of metaphysical possibilities).

According, then, to the latter, the possibility of deciding mathematical propositions which are currently undecidable relative to a background mathematical language such as ZFC should be two-dimensional. The epistemic possibility of deciding Orey sentences can thus be a guide to the metaphysical possibility thereof.[31]

The convergence between Wright's and my approaches consists, then, in that – on both approaches – there is a set of epistemically possible worlds. In the former case, the epistemically possible worlds subserve the preference rankings for the definability of expected epistemic utility. Epistemic mathematical modality is thus constitutive of the notion of rational entitlement to which Hale and Wright appeal, and – in virtue of its convergence with the two-dimensional intensional semantics here proffered – epistemically possible worlds can serve as a guide to the metaphysical mathematical possibility that mathematical propositions, such as abstraction principles for cardinals, reals, and sets, are true.

9.4   Concluding Remarks

In this essay, I have endeavored to provide an account of the modal foundations of mathematical platonism. Hale and Wright's objections to the idea that Necessitism cannot account for how possibility and actuality might converge were shown to be readily answered. In response, further, to Hale and Wright's objections to the role of epistemic and metaphysical modalities in countenancing the truth of abstraction principles and the success of mathematical predicate reference, I demonstrated how my two-dimensional modal approach to the epistemology of mathematics, augmented with Necessitism, is consistent with Hale and Wright's conception of the epistemic

---

[31]See Kanamori (2008) and Woodin (2010), for further discussion of the mathematical properties at issue.

entitlement rationally to trust that abstraction principles are true. Epistemic and metaphysical modality may thus be shown to play a constitutive role in vindicating the reality of mathematical objects and truth, and in explaining our possible knowledge thereof.

Chapter 10

Ω-Logicism: Automata, Neo-Logicism, and Set-theoretic Realism

This essay examines the philosophical significance of the consequence relation defined in the Ω-logic for set-theoretic languages. I argue that, as with second-order logic, the modal profile of validity in Ω-Logic enables the property to be epistemically tractable. Because of the dual isomorphism between algebras and coalgebras, Boolean-valued models of set theory can be interpreted as coalgebras. The modal profile of Ω-logical validity can then be countenanced within a coalgebraic logic, and Ω-logical validity can further be defined via automata (Section **2**). The philosophical significance of the foregoing is two-fold: First, that it vindicates a type of neo-logicism with regard to mathematical truth in the set-theoretic multiverse (Section **3.1**); and second, that it provides a modal and computational account of formal grasp of the concept of 'set', adducing in favor of a realist conception of the cumulative hierarchy of sets (Section **3.2**). Section **4** provides concluding remarks.

10.1 Definitions

In this section, I define the axioms of Zermelo-Fraenkel set theory with choice. I define the mathematical properties of the large cardinal axioms to which ZFC can be adjoined, and I provide a detailed characterization of the proper-

ties of Ω-logic for ZFC. Because Boolean-valued algebraic models of Ω-logic are dually isomorphic to coalgebras, a category of coalgebraic logic is then characterized which models both modal logic and deterministic automata. Modal coalgebraic models of automata are then argued to provide a precise characterization of the modal and computational profiles of Ω-logical validity.

10.1.1 Axioms[1]

- Empty set:

  $\exists x \forall u (u \notin x)$

- Extensionality:

  $x = y \iff \forall u(u \in x \iff u \in y)$

- Pairing:

  $\exists x \forall u (u \in x \iff u = a \lor u = b)$

- Union:

  $\exists x \forall u [u \in x \iff \exists v(u \in v \land v \in a)]$

- Separation:

  $\exists x \forall u [u \in x \iff u \in a \land \phi(u)]$

- Power Set:

  $\exists x \forall u (u \in x \iff u \subseteq a)$

- Infinity:

  $\exists x \varnothing \in x \land \forall u(u \in x \to \{u\} \in x)$

- Replacement:

  $\forall u \exists! v \psi(u,v) \to \forall x \exists y (\forall u \in x)(\exists v \in y)\psi(u,v)$

- Choice:

  $\forall u[u \in a \to \exists v(v \in u)] \land \forall u, x[u \in a \land x \in a \to \exists v(v \in u \iff v \in x) \lor \neg v(v \in u \land v \in x)] \to \exists x \forall u[u \in a \to \exists! v(v \in u \land u \in x)]$

---

[1]For a standard presentation, see Jech (2003). For detailed, historical discussion, see Maddy (1988,a).

10.1.2 Large Cardinals

Borel sets of reals are subsets of ω^ω or $\mathbb{R}$, closed under countable intersections and unions.[2] For all ordinals, a, such that $0 < a < \omega_1$, and $b < a$, Σ^0_a denotes the open subsets of ω^ω formed under countable unions of sets in Π^0_b, and Π^0_a denotes the closed subsets of ω^ω formed under countable intersections of Σ^0_b.

Projective sets of reals are subsets of ω^ω, formed by complementations $(\omega^\omega - u,$ for $u \subseteq \omega^\omega)$ and projections $[p(u) = \{\langle x_1, \ldots, x_n \rangle \in \omega^\omega \mid \exists y \langle x_1, \ldots, x_n, y \rangle \in u\}]$. For all ordinals a, such that $0 < a < \omega$, Π^1_0 denotes closed subsets of ω^ω; Π^1_a is formed by taking complements of the open subsets of ω^ω, Σ^1_a; and Σ^1_{a+1} is formed by taking projections of sets in Π^1_a.

The full power set operation defines the cumulative hierarchy of sets, V, such that $V_0 = \varnothing$; $V_{a+1} = P(V_0)$; and $V_\lambda = \bigcup_{a<\lambda} V_a$.

In the inner model program (cf. Woodin, 2001, 2010, 2011; Kanamori, 2012), the definable power set operation defines the constructible universe, $L(\mathbb{R})$, in the universe of sets V, where the sets are transitive such that $a \in C \iff a \subseteq C$; $L(\mathbb{R}) = V_{\omega+1}$; $L_{a+1}(\mathbb{R}) = \text{Def}(L_a(\mathbb{R}))$; and $L_\lambda(\mathbb{R}) = \bigcup_{a<\lambda}(L_a(\mathbb{R}))$.

Via inner models, Gödel (1940) proves the consistency of the generalized continuum hypothesis, $\aleph_a^{\aleph_a} = \aleph_{a+1}$, as well as the axiom of choice, relative to the axioms of ZFC. However, for a countable transitive set of ordinals, M, in a model of ZF without choice, one can define a generic set, G, such that, for all formulas, ϕ, either ϕ or $\neg\phi$ is forced by a condition, f, in G. Let $M[G] = \bigcup_{a<\kappa} M_a[G]$, such that $M_0[G] = \{G\}$; with $\lambda < \kappa$, $M_\lambda[G] = \bigcup_{a<\lambda} M_a[G]$; and $M_{a+1}[G] = V_a \cap M_a[G]$.[3] G is a Cohen real over M, and comprises a set-forcing extension of M. The relation of set-forcing, $\Vdash$, can then be defined in the ground model, M, such that the forcing condition, f, is a function from a finite subset of ω into $\{0,1\}$, and $f \Vdash u \in G$ if $f(u) = 1$ and $f \Vdash u \notin G$ if $f(u) = 0$. The cardinalities of an open dense ground model, M, and a generic extension, G, are identical, only if the countable chain condition (c.c.c.) is satisfied, such that, given a chain – i.e., a linearly ordered subset of a partially ordered (reflexive, antisymmetric, transitive) set – there is a countable, maximal antichain consisting of pairwise incompatible forcing conditions. Via set-forcing extensions, Cohen (1963, 1964) constructs a model of ZF which negates the generalized continuum hypothesis, and thus

---

[2] See Koellner (2013), for the presentation, and for further discussion, of the definitions in this and the subsequent paragraph.

[3] See Kanamori (2012,a: 2.1; 2012,b: 4.1), for further discussion.

proves the independence thereof relative to the axioms of ZF.[4]

Gödel (1946/1990: 1-2) proposes that the value of Orey sentences such as the GCH might yet be decidable, if one avails of stronger theories to which new axioms of infinity – i.e., large cardinal axioms – are adjoined.[5] He writes that: 'In set theory, e.g., the successive extensions can be represented by stronger and stronger axioms of infinity. It is certainly impossible to give a combinatorial and decidable characterization of what an axiom of infinity is; but there might exist, e.g., a characterization of the following sort: An axiom of infinity is a proposition which has a certain (decidable) formal structure and which in addition is true. Such a concept of demonstrability might have the required closure property, i.e. the following could be true: Any proof for a set-theoretic theorem in the next higher system above set theory ... is replaceable by a proof from such an axiom of infinity. It is not impossible that for such a concept of demonstrability some completeness theorem would hold which would say that every proposition expressible in set theory is decidable from present axioms plus some true assertion about the largeness of the universe of sets'.

For cardinals, x,a,C, $C \subseteq a$ is closed unbounded in a, if it is closed [if x < C and $\bigcup(C \cap a) = a$, then $a \in C$] and unbounded ($\bigcup C = a$) (Kanamori, op. cit.: 360). A cardinal, S, is stationary in a, if, for any closed unbounded $C \subseteq a$, $C \cap S \neq \emptyset$ (op. cit.). An ideal is a subset of a set closed under countable unions, whereas filters are subsets closed under countable intersections. A cardinal κ is regular if the cofinality of κ – comprised of the unions of sets with cardinality less than κ – is identical to κ. Uncountable regular limit cardinals are weakly inaccessible (op. cit.). A strongly inaccessible cardinal is regular and has a strong limit, such that if $\lambda < \kappa$, then $2^\lambda < \kappa$ (op. cit.).

Large cardinal axioms are defined by elementary embeddings.[6] Elementary embeddings can be defined thus. For models A,B, and conditions ϕ, j: A $\rightarrow$ B, $\phi\langle a_1, \ldots, a_n \rangle$ in A if and only if $\phi\langle j(a_1), \ldots, j(a_n) \rangle$ in B (363). A measurable cardinal is defined as the ordinal denoted by the critical point of j, crit(j) (Koellner and Woodin, 2010: 7). Measurable cardinals are inaccessible

---

[4]See Kanamori (2008), for further discussion.

[5]See Kanamori (2007), for further discussion. Kanamori (op. cit.: 154) notes that Gödel (1931/1986: fn48a) makes a similar appeal to higher-order languages, in his proofs of the incompleteness theorems. The incompleteness theorems are examined in further detail, in Section **3.2**, below.

[6]The definitions in the remainder of this subsection follow the presentations in Koellner and Woodin (2010) and Woodin (2010, 2011).

(Kanamori, op. cit.).

Let κ be a cardinal, and $\eta > \kappa$ an ordinal. κ is then η-strong, if there is a transitive class M and an elementary embedding, j: V $\to$ M, such that crit(j) = κ, j(κ) >η, and $V_\eta \subseteq$M (Koellner and Woodin, op. cit.).

κ is strong if and only if, for all η, it is η-strong (op. cit.).

If A is a class, κ is η-A-strong, if there is a j: V $\to$ M, such that κ is η-strong and j(A$\cap V_\kappa$)$\cap V_\eta$ = A$\cap V_\eta$ (op. cit.).

κ is a Woodin cardinal, if κ is strongly inaccessible, and for all A$\subseteq V_\kappa$, there is a cardinal $\kappa_A < \kappa$, such that κ_A is η-A-strong, for all η such that κ_η, $\eta < \kappa$ (Koellner and Woodin, op. cit.: 8).

κ is superstrong, if j: V $\to$ M, such that crit(j) = κ and $V_{j(\kappa)} \subseteq$M, which entails that there are arbitrarily large Woodin cardinals below κ (op. cit.).

Large cardinal axioms can then be defined as follows.

$\exists$xΦ is a large cardinal axiom, because:

(i) Φx is a Σ_2-formula;

(ii) if κ is a cardinal, such that V $\models \Phi(\kappa)$, then κ is strongly inaccessible; and

(iii) for all generic partial orders $\mathbb{P}\in V_\kappa$, $V^{\mathbb{P}} \models \Phi(\kappa)$; I_{NS} is a non-stationary ideal; A^G is the canonical representation of reals in L($\mathbb{R}$), i.e. the interpretation of A in M[G]; and L($\mathbb{R}$)$^{\mathbb{P}max} \models \langle$H($\omega_2$), $\in$, I_{NS}, $A^G\rangle \models$ 'ϕ'. $\mathbb{P}$ is a homogeneous partial order in L($\mathbb{R}$), such that the generic extension of L($\mathbb{R}$)$^{\mathbb{P}}$ inherits the generic invariance, i.e., the absoluteness, of L($\mathbb{R}$). Thus, L($\mathbb{R}$)$^{\mathbb{P}max}$ is (i) effectively complete, i.e. invariant under set-forcing extensions; and (ii) maximal, i.e. satisfies all Π_2-sentences and is thus consistent by set-forcing over ground models (Woodin, ms: 28).

Assume ZFC and that there is a proper class of Woodin cardinals; A$\in\mathbb{P}(\mathbb{R})$ $\cap$ L($\mathbb{R}$); ϕ is a Π_2-sentence; and V(G), s.t. $\langle$H(ω_2), $\in$, I_{NS}, $A^G\rangle \models$ 'ϕ': Then, it can be proven that L($\mathbb{R}$)$^{\mathbb{P}max} \models \langle$H($\omega_2$), $\in$, I_{NS}, $A^G\rangle \models$ 'ϕ', where 'ϕ' := $\exists$A$\in\Gamma^\infty\langle$H(ω_1), $\in$, A$\rangle \models \psi$.

The axiom of determinacy (AD) states that every set of reals, a$\subseteq\omega^\omega$ is determined, where κ is determined if it is decidable.

Woodin's (1999) Axiom (*) can be thus countenanced:

AD$^{L(\mathbb{R})}$ and L[($\mathbb{P}\omega_1$)] is a $\mathbb{P}$max-generic extension of L($\mathbb{R}$),

from which it can be derived that $2^{\aleph_0} = \aleph_2$. Thus, $\neg$CH; and so CH is absolutely decidable.

10.1.3 Ω-Logic

For partial orders, $\mathbb{P}$, let $V^{\mathbb{P}} = V^{\mathbb{B}}$, where $\mathbb{B}$ is the regular open completion of $(\mathbb{P})$.[7] $M_a = (V_a)^M$ and $M^{\mathbb{B}}_a = (V^{\mathbb{B}}_a)^M = (V_a{}^{M^{\mathbb{B}}})$. *Sent* denotes a set of sentences in a first-order language of set theory. $T\cup\{\phi\}$ is a set of sentences extending ZFC. *c.t.m* abbreviates the notion of a countable transitive $\in$-model. *c.B.a.* abbreviates the notion of a complete Boolean algebra.

Define a *c.B.a.* in V, such that $V^{\mathbb{B}}$. Let $V^{\mathbb{B}}_0 = \emptyset$; $V^{\mathbb{B}}_\lambda = \bigcup_{b<\lambda}V^{\mathbb{B}}_b$, with λ a limit ordinal; $V^{\mathbb{B}}_{a+1} = \{f\colon X \to \mathbb{B} \mid X \subseteq V^{\mathbb{B}}_a\}$; and $V^{\mathbb{B}} = \bigcup_{a\in On}V^{\mathbb{B}}_a$.

ϕ is true in $V^{\mathbb{B}}$, if its Boolean-value is $1^{\mathbb{B}}$, if and only if

$V^{\mathbb{B}} \models \phi$ iff $[\![\phi]\!]^{\mathbb{B}} = 1^{\mathbb{B}}$.

Thus, for all ordinals, a, and every *c.B.a.* $\mathbb{B}$, $V^{\mathbb{B}}_a \equiv (V_a)^{V^{\mathbb{B}}}$ iff for all $x\in V^{\mathbb{B}}$, $\exists y\in V^{\mathbb{B}}[\![x = y]\!]^{\mathbb{B}} = 1^{\mathbb{B}}$ iff $[\![x\in V^{\mathbb{B}}]\!]^{\mathbb{B}} = 1^{\mathbb{B}}$.

Then, $V^{\mathbb{B}}_a \models \phi$ iff $V^{\mathbb{B}} \models$ '$V_a \models \phi$'.

Ω-logical validity can then be defined as follows:

For $T\cup\{\phi\}\subseteq Sent$,

$T \models_\Omega \phi$, if for all ordinals a and *c.B.a.* $\mathbb{B}$, if $V^{\mathbb{B}}_a \models T$, then $V^{\mathbb{B}}_a \models \phi$.

Supposing that there exists a proper class of Woodin cardinals and if $T\cup\{\phi\}\subseteq Sent$, then for all set-forcing conditions, $\mathbb{P}$:

$T \models_\Omega \phi$ iff $V^T \models$ '$T \models_\Omega \phi$',

where $T \models_\Omega \phi \equiv \emptyset \models$ '$T \models_\Omega \phi$'.

The Ω-Conjecture states that $V \models_\Omega \phi$ iff $V^{\mathbb{B}} \models_\Omega \phi$ (Woodin, ms). Thus, Ω-logical validity is invariant in all set-forcing extensions of ground models in the set-theoretic multiverse.

The soundness of Ω-Logic is defined by universally Baire sets of reals. For a cardinal, e, let a set A be e-universally Baire, if for all partial orders $\mathbb{P}$ of cardinality e, there exist trees, S and T on ω X λ, such that $A = p[T]$ and if $G\subseteq\mathbb{P}$ is generic, then $p[T]^G = \mathbb{R}^G - p[S]^G$ (Koellner, 2013). A is universally Baire, if it is e-universally Baire for all e (op. cit.).

Ω-Logic is sound, such that $V \vdash_\Omega \phi \to V \models_\Omega \phi$. However, the completeness of Ω-Logic has yet to be resolved.

Finally, in category theory, a category C is comprised of a class Ob(C) of objects a family of arrows for each pair of objects C(A,B) (Venema, 2007: 421). A functor from a category C to a category D, **E**: C $\to$ D, is an operation mapping objects and arrows of C to objects and arrows of D (422). An endofunctor on C is a functor, **E**: C $\to$ C (op. cit.).

---

[7]The definitions in this section follow the presentation in Bagaria et al. (2006).

A **E**-coalgebra is a pair $\mathbb{A} = (A, \mu)$, with A an object of C referred to as the carrier of $\mathbb{A}$, and μ: A $\to$ **E**(A) is an arrow in C, referred to as the transition map of $\mathbb{A}$ (390).

$\mathbb{A} = \langle A, \mu\text{: } A \to \mathbf{E}(A)\rangle$ is dually isomorphic to the category of algebras over the functor μ (417-418). If μ is a functor on categories of sets, then Boolean-algebraic models of Ω-logical validity are isomorphic to coalgebraic models.

The significance of the foregoing is that coalgebraic models may themselves be availed of in order to define modal logic and automata theory. Coalgebras provide therefore a setting in which the Boolean-valued models of set theory, the modal profile of Ω-logical validity, and automata can be interdefined. In what follows, $\mathbb{A}$ will comprise the coalgebraic model – dually isomorphic to the complete Boolean-valued algebras defined in the Ω-Logic of ZFC – in which modal similarity types and automata are definable. As a coalgebraic model of modal logic, $\mathbb{A}$ can be defined as follows (407):

For a set of formulas, Φ, let $\nabla\Phi := \square \bigvee \Phi \wedge \bigwedge \diamond\Phi$, where $\diamond\Phi$ denotes the set $\{\diamond\phi \mid \phi{\in}\Phi$ (op. cit.). Then,

$\diamond\phi \equiv \nabla\{\phi, \text{T}\}$,

$\square\phi \equiv \nabla\varnothing \vee \nabla\phi$ (op. cit.).

Let an **E**-coalgebraic modal model, $\mathbb{A} = \langle S,\lambda,R[.]\rangle$, such that $\mathbb{S},s \Vdash \nabla\Phi$ if and only if, for all (some) successors σ of s$\in$S, $[\Phi,\sigma(\text{s}){\in}\mathbf{E}(\Vdash_{\mathbb{A}})]$ (op. cit.).

A coalgebraic model of deterministic automata can finally be thus defined (391). An automaton is a tuple, $\mathbb{A} = \langle A, a_I, C, \delta, F\rangle$, such that A is the state space of the automaton $\mathbb{A}$; $a_I{\in}A$ is the automaton's initial state; C is the coding for the automaton's alphabet, mapping numerals to properties of the natural numbers; δ: A X C $\to$ A is a transition function, and F$\subseteq$ A is the collection of admissible states, where F maps A to $\{1,0\}$, such that F: A $\to$ 1 if a$\in$F and A $\to$ 0 if a$\notin$F (op. cit.). The determinacy of coalgebraic automata, the category of which is dually isomorphic to the Set category satisfying Ω-logical consequence, is secured by the existence of Woodin cardinals: Assuming ZFC, that λ is a limit of Woodin cardinals, that there is a generic, set-forcing extension G $\subseteq$ the collapse of $\omega < \lambda$, and that $\mathbb{R}^* = \bigcup\{\mathbb{R}^G[a] \mid a < \lambda\}$, then $\mathbb{R}^* \models$ the axiom of determinacy (AD) (Koellner and Woodin, op. cit.: 10).

Thus, $\mathbb{A}$ is the coalgebraic category for modal, deterministic automata, dually isomorphic to the complete Boolean-valued algebraic models of Ω-logical validity, as defined in the category of sets.

10.2   Discussion

This section examines the philosophical significance of the foregoing, and argues that, similarly to second-order logical consequence, (i) the 'mathematical entanglement' of Ω-logical validity does not undermine its status as a relation of pure logic; and (ii) both the modal profile and model-theoretic characterization of Ω-logical consequence provide a guide to its epistemic tractability.[8] I argue, finally, that there are several considerations adducing in favor of the claim that the interpretation of the concept of set constitutively involves modal notions. The role of the category of modal coalegebraic deterministic automata in (i) characterizing the modal profile of Ω-logical consequence, and (ii) being constitutive of the formal understanding-conditions for the concept of set, provides, then, support for a realist conception of the cumulative hierarchy.

10.2.1   Neo-Logicism

Frege's (1884/1980; 1893/2013) proposal – that cardinal numbers can be explained by specifying an equivalence relation, expressible in the signature of second-order logic and identity, on lower-order representatives for higher-order entities – is the first attempt to provide a foundation for mathematics on the basis of logical axioms rather than rational or empirical intuition. In Frege (1884/1980. cit.: 68) and Wright (1983: 104-105), the number of the concept, **A**, is argued to be identical to the number of the concept, **B**, if and only if there is a one-to-one correspondence between **A** and **B**, i.e., there is a bijective mapping, R, from **A** to **B**. With Nx: a numerical term-forming operator,

- $\forall \mathbf{A} \forall \mathbf{B} \exists R[[\text{Nx: } \mathbf{A} = \text{Nx: } \mathbf{B} \equiv \exists R[\forall x[\mathbf{A}x \to \exists y(\mathbf{B}y \wedge Rxy \wedge \forall z(\mathbf{B}z \wedge Rxz \to y = z))] \wedge \forall y[\mathbf{B}y \to \exists x(\mathbf{A}x \wedge Rxy \wedge \forall z(\mathbf{A}z \wedge Rzy \to x = z))]]]$.

Frege's Theorem states that the Dedekind-Peano axioms for the language of arithmetic can be derived from the foregoing abstraction principle, as augmented to the signature of second-order logic and identity.[9]   Thus, if

---

[8]The phrase, 'mathematical entanglement', is owing to Koellner (2010: 2).

[9]Cf. Dedekind (1888/1963) and Peano (1889/1967). See Wright (1983: 154-169) for a proof sketch of Frege's theorem; Boolos (1987) for the formal proof thereof; and Parsons (1964) for an incipient conjecture of the theorem's validity.

second-order logic may be counted as pure logic, despite that domains of second-order models are definable via power set operations, then one aspect of the philosophical significance of the abstractionist program consists in its provision of a foundation for classical mathematics on the basis of pure logic as augmented with non-logical implicit definitions expressed by abstraction principles.

There are at least three reasons for which a logic defined in ZFC might not undermine the status of its consequence relation as being logical. The first reason for which the mathematical entanglement of Ω-logical validity might be innocuous is that, as Shapiro (1991: 5.1.4) notes, many mathematical properties cannot be defined within first-order logic, and instead require the expressive resources of second-order logic. For example, the notion of well-foundedness cannot be expressed in a first-order framework, as evinced by considerations of compactness. Let E be a binary relation. Let m be a well-founded model, if there is no infinite sequence, $a_0, \ldots, a_i$, such that $Ea_0, \ldots,$ Ea_{i+1} are all true. If m is well-founded, then there are no infinite-descending E-chains. Suppose that T is a first-order theory containing m, and that, for all natural numbers, n, there is a T with $n + 1$ elements, $a_0, \ldots, a_n$, such that $\langle a_0, a_1 \rangle, \ldots, \langle a_n, a_{n-1} \rangle$ are in the extension of E. By compactness, there is an infinite sequence such that that $a_0 \ldots a_i$, s.t. $Ea_0, \ldots, Ea_{i+1}$ are all true. So, m is not well-founded.

By contrast, however, well-foundedness can be expressed in a second-order framework:

$\forall X[\exists x Xx \to \exists x[Xx \land \forall y(Xy \to \neg Eyx)]]$, such that m is well-founded iff every non-empty subset X has an element x, s.t. nothing in X bears E to x.

One aspect of the philosophical significance of well-foundedness is that it provides a distinctively second-order constraint on when the membership relation in a given model is intended. This contrasts with Putnam's (1980) claim, that first-order models *mod* can be intended, if every set s of reals in *mod* is such that an ω-model in *mod* contains s and is constructible, such that – given the Downward Lowenheim-Skolem theorem[10] – if *mod* is non-constructible but has a submodel satisfying 's is constructible', then the model is non-well-founded and yet must be intended. The claim depends

---

[10]For any first-order model M, M has a submodel M' whose domain is at most denumerably infinite, s.t. for all assignments s on, and formulas $\phi(x)$ in, M', $M,s \Vdash \phi(x) \iff M',s \Vdash \phi(x)$.

on the assumption that general understanding-conditions and conditions on intendedness must be co-extensive, to which I will return in Section **3.2**

The second reason for which Ω-logic's mathematical entanglement might not be pernicious is that first-order models of arithmetic and ZFC themselves include non-logical, mathematical vocabulary. For example, first-order ω-logic has domains comprised of ordinal-definable natural numbers, and – crucially – first-order β-logic has domains comprised of well-founded ordinal-definable natural numbers (Koellner, 2010: 11-12).

Finally, a third reason for which the consequence relation specified in the Ω-logic of set theory is genuinely logical may again be appreciated by its comparison with second-order logic. Shapiro (1998) defines the model-theoretic characterization of logical consequence as follows:

'(10) Φ is a logical consequence of [a model] Γ if Φ holds in all possibilities under every interpretation of the nonlogical terminology which holds in Γ' (148).

A condition on the foregoing is referred to as the 'isomorphism property', according to which 'if two models M, M' are isomorphic vis-a-vis the non-logical items in a formula Φ, then M satisfies Φ if and only if M' satisfies Φ' (151).

Shapiro argues, then, that the consequence relation specified using second-order resources is logical, because of its modal and epistemic profiles. The epistemic tractability of second-order validity consists in 'typical soundness theorems, where one shows that a given deductive system is 'truth-preserving' (154). He writes that: '[I]f we know that a model is a good mathematical model of logical consequence (10), then we know that we won't go wrong using a sound deductive system. Also, we can know that an argument is a logical consequence . . . via a set-theoretic proof in the metatheory' (154-155).

The modal profile of second-order validity provides a second means of accounting for the property's epistemic tractability. Shapiro argues, e.g., that: 'If the isomorphism property holds, then in evaluating sentences and arguments, the only 'possibility' we need to 'vary' is the size of the universe. If enough sizes are represented in the universe of models, then the modal nature of logical consequence will be registered . . . [T]he only 'modality' we keep is 'possible size', which is relegated to the set-theoretic metatheory' (152).

Shapiro's remarks about the considerations adducing in favor of the logicality of non-effective, second-order validity generalize to Ω-logical validity. In the previous section, the modal profile of Ω-logical validity was codified

by the dual isomorphism between complete Boolean-valued algebraic models of Ω-logic and the category, $\mathbb{A}$, of coalgebraic modal logics. As with Shapiro's definition of logical consequence, where Φ holds in all possibilities in the universe of models and the possibilities concern the 'possible size' in the set-theoretic metatheory, the Ω-Conjecture states that $V \models_\Omega \phi$ iff $V^{\mathbb{B}} \models_\Omega \phi$, such that Ω-logical validity is invariant in all set-forcing extensions of ground models in the set-theoretic multiverse.

Finally, the epistemic tractability of Ω-logical validity is secured, both – as on Shapiro's account of second-order logical consequence – by its soundness, but also by its isomorphism to the coalgebraic category of deterministic automata, where the determinacy thereof is again secured by the existence of Woodin cardinals.

10.2.2 Set-theoretic Realism

In this section, I argue, finally, that the modal profile of Ω-logic can be availed of in order to account for the understanding-conditions of the concept of set, and thus crucially serve as part of the argument for set-theoretic realism.

Putnam (op. cit.: 473-474) argues that defining models of first-order theories is sufficient for both understanding and specifying an intended interpretation of the latter. Wright (1985: 124-125) argues, by contrast, that understanding-conditions for mathematical concepts cannot be exhausted by the axioms for the theories thereof, even on the intended interpretations of the theories. He suggests, e.g., that:

'[I]f there really were uncountable sets, their existence would surely have to flow from the concept of set, as intuitively satisfactorily explained. Here, there is, as it seems to me, no assumption that the content of the ZF-axioms cannot exceed what is invariant under all their classical models. [Benacerraf] writes, e.g., that: 'It is granted that they are to have their 'intended interpretation': 'e' is to mean set-membership. Even so, and conceived as encoding the intuitive concept of set, they fail to entail the existence of uncountable sets. So how can it be true that there are such sets? Benacerraf's reply is that the ZF-axioms are indeed faithful to the relevant informal notions only if, in addition to ensuring that 'E' means set-membership, we interpret them so as to observe the constraint that 'the universal quantifier has to mean all or at least all sets' (p. 103). It follows, of course, that if the concept of set does determine a background against which Cantor's theorem, under its intended interpretation, is sound, there is more to the concept of

set that can be explained by communication of the intended sense of 'e' and the stipulation that the ZF-axioms are to hold. And the residue is contained, presumably, in the informal explanations to which, Benacerraf reminds us, Zermelo intended his formalization to answer. At least, this must be so if the 'intuitive concept of set' is capable of being explained at all. Yet it is notable that Benacerraf nowhere ventures to supply the missing informal explanation – the story which will pack enough into the extension of 'all sets' to yield Cantor's theorem, under its intended interpretation, as a highly non-trivial corollary'(op. cit).

In order to provide the foregoing explanation in virtue of which the concept of set can be shown to be associated with a realistic notion of the cumulative hierarchy, I will argue that there are several points in the model theory and epistemology of set-theoretic languages at which the interpretation of the concept of set constitutively involves modal notions. The aim of the section will thus be to provide a modal foundation for mathematical platonism.

One point is in the coding of the signature of the theory, T, in which Gödel's incompleteness theorems are proved (cf. Halbach and Visser, 2014). Relative to,

(i) a choice of coding for an ω-complete, recursively axiomatizable language, L, of T – i.e. a mapping between properties of numbers and properties of terms and formulas in L;

(ii) a predicate, phi; and

(iii) a fixed-point construction:

Let phi express the property of 'being provable', and define (iii) such that, for all consistent theories T of L, there are sentences, p_{phi}, corresponding to each formula, phi(x), in T, s.t. for 'm' := p_{phi},

$\vdash_T p_{phi}$ iff phi(m).

One can then construct a sentence, 'm' := ¬phi(m), such that L is incomplete (the first incompleteness theorem).

Moreover, L cannot prove its own consistency:

If:

$\vdash_T$ 'm' iff ¬phi(m),

Then:

$\vdash_T C \rightarrow m$.

Thus, L is consistent only if L is inconsistent (the second incompleteness theorem).

162

In the foregoing, the choice of coding bridges the numerals in the language with the properties of the target numbers. The choice of coding is therefore intensional, and has been marshalled in order to argue that the very notion of syntactic computability – via the equivalence class of partial recursive functions, λ-definable terms, and the transition functions of discrete-state automata such as Turing machines – is constitutively semantic (cf. Rescorla, 2015). Further points at which intensionality can be witnessed in the phenomenon of self-reference in arithmetic are introduced by Reinhardt (1986). Reinhardt (op. cit.: 470-472) argues that the provability predicate can be defined relative to the minds of particular agents – similarly to Quine's (1968) and Lewis' (1979) suggestion that possible worlds can be centered by defining them relative to parameters ranging over tuples of spacetime coordinates or agents and locations – and that a theoretical identity statement can be established for the concept of the foregoing minds and the concept of a computable system.

In the previous section, intensional computational properties were defined via modal coalgebraic deterministic automata, where the coalgebraic categories are dually isomorphic to the category of sets in which Ω-logical validity was defined. Coalgebraic modal logic was shown to elucidate the modal profile of Ω-logical consequence in the Boolean-valued algebraic models of set theory. The intensionality witnessed by the choice of coding may therefore be further witnessed by the modal automata specified in the foregoing coalgebraic logic.

A second point at which understanding-conditions may be shown to be constitutively modal can be witnessed by the conditions on the epistemic entitlement to assume that the language in which Gödel's second incompleteness theorem is proved is consistent (cf. Dummett, 19; Wright, 1985). Wright (op. cit.: 91, fn.9) suggests that '[T]o treat [a] proof as establishing consistency is implicitly to exclude any doubt ... about the consistency of first-order number theory'. Wright's elaboration of the notion of epistemic entitlement, appeals to a notion of rational 'trust', which he argues is recorded by the calculation of 'expected epistemic utility' in the setting of decision theory (2004; 2014: 226, 241). Wright notes that the rational trust subserving epistemic entitlement will be pragmatic, and makes the intriguing point that 'pragmatic reasons are not a special genre of reason, to be contrasted with e.g. epistemic, prudential, and moral reasons' (2012: 484). Crucially, however, the very idea of expected epistemic utility in the setting of decision theory makes implicit appeal to the notion of possible worlds, where the

latter can again be determined by the coalgebraic logic for modal automata.

A third consideration adducing in favor of the thought that grasp of the concept of set might constitutively possess a modal profile is that the concept can be defined as an intension – i.e., a function from possible worlds to extensions. The modal similarity types in the coalgebraic modal logic may then be interpreted as dynamic-interpretational modalities, where the dynamic-interpretational modal operator has been argued to entrain the possible reinterpretations both of the domains of the theory's quantifiers (cf. Fine, 2005, 2006), as well as of the extensions of terms expressing the membership relation (cf. Uzquiano, 2015,a).

The fourth consideration avails directly of the modal profile of Ω-logical consequence. While the above dynamic-interpretational modality will suffice for possible reinterpretations of mathematical terms, the absoluteness and generic invariance of the consequence relation is such that, if the Ω-conjecture is true, then Ω-logical validity is invariant in all possible set-forcing extensions of ground models in the set-theoretic multiverse. The truth of the Ω-conjecture would thereby place an indefeasible necessary condition on a formal understanding of the intension for the concept of set.

10.3   Concluding Remarks

In this essay I have examined the philosophical significance of the isomorphism between Boolean-valued algebraic models of modal Ω-logic and modal coalgebraic models of automata. I argued that – as with the property of validity in second-order logic – Ω-logical validity is genuinely logical, and thus entails a type of neo-logicism in the foundations of mathematics. I argued, then, that modal coalegebraic deterministic automata, which characterize the modal profile of Ω-logical consequence, are constitutive of the interpretation of mathematical concepts such as the membership relation. The philosophical significance of modal Ω-logic is thus that it can be availed of to vindicate both a neo-logicist foundation for set theory and a realist interpretation of the cumulative hierarchy of sets.

Chapter 11

Epistemic Modality and Absolute Decidability

This essay aims to contribute to the analysis of the nature of mathematical modality, and to the applications of the latter to unrestricted quantification and absolute decidability. I argue that mathematical modality falls under at least three types; the interpretational, the metaphysical, and the logical. The interpretational type of mathematical modality has traditionally been taken to concern the interpretation of the quantifiers (cf. Linnebo, 2009, 2010, 2013; Studd, 2013); the possible reinterpretations of the intensions of the concept of set (Uzquiano, 2015,a); and the possibility of reinterpreting the domain over which the quantifiers range, in order to avoid inconsistency (cf. Fine, 2005, 2006, 2007). The metaphysical type of modality concerns the ontological profile of abstracta and mathematical truth. Abstracta are thus argued to have metaphysically necessary being, and mathematical truths hold of metaphysical necessity, if at all (cf. Fine, 1981; Williamson, 2016). Instances, finally, of the logical type of mathematical modality might concern the properties of consistency (cf. Field, 1989: 249-250, 257-260; Rayo, 2013: 50; Leng: 2007; 2010: 258), and can perhaps be further witnessed by the logic of provability (cf. Boolos, 1993) and the modal profile of forcing (cf. Kripke 1965; Hamkins and Löwe, 2008).

The significance of the present contribution is as follows. (i) Rather than countenancing the interpretational type of mathematical modality as a primitive, I argue that the interpretational type of mathematical modality is a

species of epistemic modality.[1] (ii) I argue, then, that the framework of two-dimensional intensional semantics ought to be applied to the mathematical setting. The framework permits of a formally precise account of the priority and relation between epistemic mathematical modality and metaphysical mathematical modality. I target, in particular, the modal axioms that the respective interpretations of the modal operator ought to satisfy. The discrepancy between the modal systems governing the parameters in the two-dimensional intensional setting provides an explanation of the difference between the metaphysical possibility of absolute decidability and our knowledge thereof. (iii) Finally, I examine the application of the mathematical modalities beyond the issues of unrestricted quantification and indefinite extensibility. As a test case for the two-dimensional approach, I investigate the interaction between the epistemic and metaphysical mathematical modalities and large cardinal axioms. The two-dimensional intensional framework permits of a formally precise means of demonstrating how the metaphysical possibility of absolute decidability and the continuum hypothesis can be accessed by their epistemic-modal-mathematical profile. The logical mathematical modalities – of consistency, provability, and forcing – provide the means for discerning whether mathematical truths are themselves epistemically possible. I argue that, in the absence of disproof, large cardinal axioms are epistemically possible, and thereby provide a sufficient guide to the metaphysical mathematical possibility of determinacy claims and the continuum hypothesis.

In Section **2**, I define the formal clauses and modal axioms governing the epistemic and metaphysical types of mathematical modality. In Section **3**, I discuss how the properties of the epistemic mathematical modality and metaphysical mathematical modality converge and depart from previous at-

---

[1]A precedent to the current approach is Parsons (1979-1980; 1983: p. 25, chs.10-11; 2008: 176), who argues that intuition is both a species of the imagination and can be formalized by a mathematical modality. The mathematical modality is governed by S4.2, and concerns possible iterations of the successor operation in arithmetic and possible extensions of the set-theoretic cumulative hierarchy. Among the differences between Parsons' approach and the one here outlined is (i) that, by contrast to the current proposal, Parsons notes that his notion of mathematical modality is not epistemic (2008: 81fn1); and (ii) that Parsons (1997: 348-351; 2008: 98-100) suggests that the intuitional mathematical modality concerning computable functions is an idealization insensitive to distinctions such as those captured by computational complexity theory, rather than being defined relative to an epistemic modal space comprising the computational theory of mind. (See Chapters **2-3**, for further discussion.)

tempts to delineate the contours of similar notions. Section **4** extends the two-dimensional intensional framework to the issue of mathematical knowledge; in particular, to the modal profile of large cardinal axioms and to the absolute decidability of the continuum hypothesis. Section **5** provides concluding remarks.

11.1 Mathematical Modality

11.1.1 Metaphysical Mathematical Modality

A formula is a logical truth if and only if the formula is true in an intended model structure, M = <W, D, R, V>, where W designates a space of metaphysically possible worlds; D designates a domain of entities, constant across worlds; R designates an accessibility relation on worlds; and V is an assignment function mapping elements in D to subsets of W. A formula in M is a modal truth if and only if a faithful interpretation maps the formula to a metaphysically universal proposition (cf. Williamson, 2013.: 106). A formula satisfies conditions on metaphysically universality if and only if the formula is true on its universal generalization (cf. Williamson, op. cit.: 93).

11.1.2 Epistemic Mathematical Modality

In order to accommodate the notion of epistemic possibility, we enrich M with the following conditions: M = <C, W, D, R, V>, where C, a set of epistemically possibilities, is constrained as follows:

Let $[\![\phi]\!]^c \subseteq$ C;

(ϕ is a formula encoding a state of information at an epistemically possible world).

-$\texttt{pri}(x) = \lambda c.[\![x]\!]^{c,c}$;

(the two parameters relative to which x – a propositional variable – obtains its value are epistemically possible worlds. The function from possible formulas to values is thus an intension).

-$\texttt{sec}(x) = \lambda c.[\![x]\!]^{w,w}$

(the two parameters relative to which x obtains its value are metaphysically possible worlds).

Then:

- **Epistemic Mathematical Necessity (Apriority)**

$[\![\blacksquare\phi]\!]^{c,w} = 1 \iff \forall c'[\![\phi]\!]^{c,c'} = 1$

(ϕ is true at all points in epistemic modal space).

- **Epistemic Mathematical Possibility**

  $[\![\diamond_c\phi]\!]^{i_n} \neq \emptyset \iff [\![Pr\phi]\!]^{i_n} \neq \emptyset \land >.5$, else $\langle\emptyset, Pr_{i_n}(\phi \mid \emptyset)\rangle$, where i_n designates an agent's state of information i in a context n.

  (ϕ might be true if and only if its value is not null and it is greater than .5).

Crucially, epistemic mathematical modality is constrained by consistency, and the formal techniques of provability and forcing. A mathematical formula is false, and therefore metaphysically impossible, if it can be disproved or induces inconsistency in a model.

11.1.3   Interaction

- **Convergence**

  $\forall c \exists w [\![\phi]\!]^{c,w} = 1$

  (the value of x is relative to a parameter for the space of epistemically possible worlds. The value of x relative to the first parameter determines the value of x relative to the second parameter for the space of metaphysical possibility).

- **Super-rigidity (2D-Intension):**

  $[\![\phi]\!]^{w,c} = 1 \iff \forall w',c'[\![\phi]\!]^{w',c'} = 1$

  (the intension of ϕ is rigid in all points in metaphysical and epistemic modal space).

11.1.4   Modal Axioms

- Metaphysical mathematical modality is governed by the modal system KTE, as augmented by the Barcan formula and its Converse (cf. Fine, 1981).

K: $\Box[\phi \to \psi] \to [\Box\phi \to \Box\psi]$
T: $\Box\phi \to \phi$

E: $\neg\Box\phi \rightarrow \Box\neg\Box\phi$
Barcan: $\diamond\exists x Fx \rightarrow \exists x \diamond Fx$
Converse Barcan: $\exists x \diamond Fx \rightarrow \diamond\exists x Fx$

- Epistemic mathematical modality is governed by the modal system, KT4, as augmented by the Barcan formula and the Converse Barcan formula.[2]

K: $\blacksquare[\phi \rightarrow \psi] \rightarrow [\blacksquare\phi \rightarrow \blacksquare\psi]$
T: $\blacksquare\phi \rightarrow \phi$
4: $\blacksquare\phi \rightarrow \blacksquare\blacksquare\phi$
Barcan: $\blacklozenge\exists x Fx \rightarrow \exists x \blacklozenge Fx$
Converse Barcan: $\exists x \blacklozenge Fx \rightarrow \blacklozenge\exists x Fx$

11.2   Departures from Precedent

The approach to mathematical modality, according to which it yields a representation of the cumulative universe of sets, has been examined by Fine (2005; 2006) and Uzquiano (op. cit.). Fine argues that the mathematical modality should be interpretational; and thus taken to concern the reinterpretation of the domain over which the quantifiers range, in order to avoid inconsistency. Uzquiano argues similarly for an interpretational construal of mathematical modality, where the cumulative hierarchy of sets is fixed, yet

---

[2]Reasons adducing against including the Smiley-Gödel-Löb provability formula among the axioms of epistemic mathematical modality are examined in Section 5. GL states that '$\blacksquare[\blacksquare\phi \rightarrow \phi] \rightarrow \blacksquare\phi$'. For further discussion of the properties of GL, see Löb (1955); Smiley (1963); Kripke (1965); and Boolos (1993). Löb's provability formula was formulated in response to Henkin's (1952) problem concerning whether a sentence which ascribes the property of being provable to itself is provable. (Cf. Halbach and Visser, 2014, for further discussion.) For an anticipation of the provability formula, see Wittgenstein (1933-1937/2005: 378). Wittgenstein writes: 'If we prove that a problem can be solved, the concept 'solution' must somehow occur in the proof. (There must be something in the mechanism of the proof that corresponds to this concept.) But the concept mustn't be represented by an external description; it must really be demonstrated. / The proof of the provability of a proposition is the proof of the proposition itself' (op. cit.). Wittgenstein contrasts the foregoing type of proof with 'proofs of relevance' which are akin to the mathematical, rather than empirical, propositions, discussed in Wittgenstein (2001: IV, 4-13, 30-31).

what is possibly reinterpreted is the non-logical vocabulary of the language, in particular the membership relation.[3]

On Fine's approach, the interpretational modality is both postulational, and 'prescriptive' or imperatival. The prescriptive element consists in the rule:

'Introduction: !x.C(x)',

such that one is enjoined to postulate, i.e. to 'introduce an object x conforming to the condition C(x)' (2005: 91; 2006: 38).

In the setting of unrestricted quantification, suppose, e.g., that there is an interpretation for the domain over which a quantifier ranges. Fine writes that an interpretation 'I is exten[s]ible – in symbols, $E(I)$ – if possibly some interpretation extends it, i.e. $\diamond \exists J(I \subset J)$' (2006: 30). Then, the interpretation of the domain over which the quantifier ranges is *extensible*, if '$\forall I.E(I)$'. The interpretation of the domain over which the quantifier ranges is *indefinitely extensible*, if '$\Box \forall I.E(I)$' iff '$\Box \forall I \diamond \exists J(I \subset J)$', where the reinterpretation is induced via the prescriptive imperative to postulate the existence of a new object by the foregoing 'Introduction' rule (2006: 30-31; 38). Fine clarifies that the interpretational approach is consistent with a 'realist ontology' of the set of reals. He refers to the imperative to postulate new objects, and thereby reinterpret the domain for the quantifier, as the 'mechanism' by which epistemically to track the cumulative hierarchy of sets (2007: 124-125).

In accord with Fine's approach, the epistemic mathematical modality defined in the previous section was taken to have a similarly representational interpretation, and perhaps the postulational property is an optimal means of inducing a reinterpretation of the domain of the quantifier. However, the present approach avoids a potential issue with Fine's account, with regard to the the introduction of deontic modal properties of the prescriptive and imperatival rules that he mentions.[4] It is sufficient that the interpretational modalities are a species of epistemic modality, i.e. possibilities that are relative to agents' spaces of states of information.

Developing Fine's program, Linnebo (2013) outlines a modalized version of ZF. Similarly to the modal axioms for the epistemic mathematical modality specified in the previous section, Linnebo argues that his modal set theory ought to be governed by the system S4.2, the Converse Barcan formula, and

---

[3]Compare Gödel, 1947; Williamson, 1998; and Fine, 2005.

[4]For an analysis of the precise interaction between the semantic values of epistemic and deontic modal operators, see Author (ms).

(at least a restricted version of) the Barcan formula. However – rather than being either interpretational or epistemic – Linnebo deploys the mathematical modality in order to account for the notion of 'potential infinity', as anticipated by Aristotle.[5] The mathematical modality is thereby intended to provide a formally precise answer to the inquiry into the extent of the cumulative set-theoretic hierarchy; i.e., in order to precisify the answer that the hierarchy extends 'as far as possible' (2013: 205).[6]

Thus, Linnebo takes the modality to be constitutive of the actual ontology of sets; and the quantifiers ranging over the actual ontology of sets are claimed to have an 'implicitly modal' profile (2010: 146; 2013: 225). He suggests, e.g., that: 'As science progresses, we formulate set theories that characterize larger and larger initial segments of the universe of sets. At any one time, precisely those sets are actual whose existence follows from our strongest, well-established set theory' (2010: 159n21). However – despite his claim that the modality is constitutive of the actual ontology of sets – Linnebo concedes that the mathematical modality at issue cannot be interpreted metaphysically, because sets exist of metaphysical necessity if at all (2010: 158; 2013: 207). In order partly to allay the tension, Linnebo remarks, then, that set theorists 'do not regard themselves as located at some particular stage of the process of forming sets' (2010: 159); and this might provide evidence that the inquiry – concerning at which stage in the process of set-individuation we happen to be, at present – can be avoided.

Another distinction to note is that both Linnebo (op. cit.) and Uzquiano (op. cit.) avail of second-order plural quantification, in developing their primitivist and interpretational accounts of mathematical modality. By contrast to their approaches, the epistemic and metaphysical modalities defined in the previous section are defined with second-order singular quantification over sets.

Finally, Linnebo and Uzquiano both suggest that their mathematical

---

[5]Cf. Aristotle, *Physics*, Book III, Ch. 6.

[6]Precursors to the view that modal operators can be availed of in order to countenance the potential hierarchy of sets include Hodes (1984,a). Intensional constructions of set theory are further developed by Reinhardt (1974); Parsons (op. cit.); Myhill (1985); Scedrov (1985); Flagg (1985); Goodman (1985); Hellman (1990); Nolan (2002); and Studd (2013). (See Shapiro (1985) for an intensional construction of arithmetic.) Parsons (19Chihara (2004: 171-198) argues that 'broadly logical' conceptual possibilities can be used to represent imaginary situations relevant to the construction of open-sentence tokens. The open-sentences can then be used to define the properties of natural and cardinal numbers and the axioms of Peano arithmetic.

modalities ought to be governed by the G axiom; i.e. $\Diamond\Box\phi \rightarrow \Box\Diamond\phi$. The present approach eschews, however, of the G axiom, in virtue of the following. Williamson (2009) demonstrates that – because KT4G is a sublogic of S5 – an epistemic operator which validates the conjunction of the 4 axiom of positive introspection and the E axiom of negative introspection will be inconsistent with the condition of 'recursively enumerable quasi-conservativeness'. Recursively enumerable quasi-conservativeness is a computational constraint on an epistemic agent's theorizing, according to which the intended models of the agent's theory are both maximally consistent and conservatively extended by addition of the 'box'-operator, interpreted as expressing the agent's state of knowledge. As axioms of an agent's consistent, recursively axiomatizable theorizing about the theory of its own states of knowledge and belief, the conjunction of 4 and E would entail that the agent's theory is both consistent and decidable, in conflict with Gödel's (1931) second incompleteness theorem. The modal system, KT4, avoids the foregoing result. In the present setting, the circumvention is innocuous, because the undecidability – yet recursively enumerable quasi-conservativeness – of an epistemic agent's consistent theorizing about its epistemic states is consistent with the epistemic mathematical possibility that large cardinal axioms are absolutely decidable.

11.3   Knowledge of Absolute Decidability

Williamson (2016) examines the extension of the metaphysically modal profile of mathematical truths to the question of absolute decidability. In this section, I aim to extend Williamson's analysis to the notion of epistemic mathematical modality that has been developed in the foregoing sections. The extension provides a crucial means of witnessing the significance of the two-dimensional intensional approach for the epistemology of mathematics.

Williamson proceeds by suggesting the following line of thought. Suppose that A is a true interpreted mathematical formula which eludes present human techniques of provability; e.g. the continuum hypothesis (op. cit.). Williamson argues that mathematical truths are metaphysically necessary (op. cit.). From there, he suggests that knowledge of A satisfies the condition of *safety* from error, as codified via a reflexive and symmetric accessibility relation from worlds at which A is known. Thus, there is either no, or a small risk of, not believing that A, relative to a world in which A is known –

although the safety condition is not itself sufficient for mathematical knowledge that A. Williamson then enjoins one to consider the following scenario: It is metaphysically possible that there is a species which can prove that A. Therefore, A is absolutely provable; that is, A 'can in principle be known by a normal mathematical process' such as derivation in an axiomatizable formal system with quantification and identity.

Williamson's scenario evinces one issue for the 'back-tracking' approach to modal epistemology, at least as it might be applied to the issue of possible mathematical knowledge. On the back-tracking approach, the method of modal epistemology is taken to proceed by first discerning the metaphysical modal truths – normally by natural-scientific means – and then working backward to the exigent incompleteness of an individual's epistemic states concerning such truths (cf. Stalnaker, 2003; Vetter, 2013).

The issue for the back-tracking method that Williamson's scenario illuminates is that the metaphysical mathematical possibility that CH is absolutely decidable must in some way converge with the epistemic possibility thereof. However, the normal mathematical techniques that Williamson specifies – i.e. proof and forcing – fall within the remit of what is mathematically possible relative to agents' states of information; i.e. what is epistemically mathematically possible. Thus, whether CH is metaphysically necessary – and thus, as Williamson claims, metaphysically possible and absolutely decidable thereby – can only be witnessed by the epistemic means of demonstrating that its absolute decidability is not impossible. It may thus be epistemically possible that Williamson's technically advanced species, which can absolutely decide CH, exist – following Williamson (2013), they actually exist, albeit non-concretely – but the metaphysical necessity of the absolute decidability of CH needs still to be corroborated.

The significance of the two-dimensional intensional framework outlined in the foregoing is that it provides an explanation of the discrepancy between metaphysical mathematical modality and epistemic mathematical modality. The metaphysical mathematical modality is taken to be more fundamental than the epistemic, as witnessed by the order of the parameters specified in the Convergence property in Section **2**. Further and crucially, metaphysical mathematical modality is governed by the system S5, the Barcan formula, and its Converse, whereas epistemic mathematical modality is governed by KT4, the Barcan formula, and its Converse. Thus, epistemic mathematical modality figures as the mechanism, which enables the tracking of metaphys-

173

ically possible mathematical truth.[7]

Leitgeb (2009) endeavors similarly to argue for the convergence between the notion of informal provability – countenanced as an epistemic modal operator, K – and mathematical truth. Availing of Hilbert's (1923/1996: ¶18-42) epsilon terms for propositions, such that, for an arbitrary predicate, $\mathbf{C}(x)$, with x a propositional variable, the term '$\epsilon p.\mathbf{C}(p)$' is intuitively interpreted as stating that 'there is a proposition, x(/p), s.t. the formula, that p satisfies $\mathbf{C}$, obtains' (op. cit.: 290). Leitgeb purports to demonstrate that $\forall p(p \rightarrow Kp)$, i.e. that informal provability is absolute; i.e. truth and provability are co-extensive. He argues as follows. Let A(p) abbreviate the formula 'p $\wedge$ $\neg K(p)$', i.e., that the proposition, p, is true while yet being unprovable. Let K be the informal provability operator reflecting knowability or epistemic necessity, with $\langle K \rangle$ its dual.[8] Then:

1. $\exists p(p \wedge \neg Kp) \iff \epsilon p.A(p) \wedge \neg K\epsilon p.A(p)$.

---

[7]A provisional definition of large cardinal axioms is as follows.

$\exists x \Phi$ is a large cardinal axiom, because:

(i) Φx is a Σ_2-formula;

(ii) if κ is a cardinal, such that $V \models \Phi(\kappa)$, then κ is strongly inaccessible, where a cardinal κ is regular if the cofinality of κ – comprised of the unions of sets with cardinality less than κ – is identical to κ, and a strongly inaccessible cardinal is regular and has a strong limit, such that if $\lambda < \kappa$, then $2^\lambda < \kappa$ (Cf. Kanamori, 2012: 360); and

(iii) for all generic partial orders $\mathbb{P} \in V_\kappa$, $V^\mathbb{P} \models \Phi(\kappa)$; I_{NS} is a non-stationary ideal, where an ideal is a subset of a set closed under countable unions, whereas filters are subsets closed under countable intersections. (Cf. Kanamori, op. cit.: 361); A^G is the canonical representation of reals in $L(\mathbb{R})$, i.e. the interpretation of A in M[G]; $H(\kappa)$ is comprised of all of the sets whose transitive closure is $< \kappa$ (cf. Rittberg, 2015); and $L(\mathbb{R})^{\mathbb{P}max} \models \langle H(\omega_2), \in, I_{NS}, A^G \rangle \models$ 'ϕ'. $\mathbb{P}$ is a homogeneous partial order in $L(\mathbb{R})$, such that the generic extension of $L(\mathbb{R})^\mathbb{P}$ inherits the generic invariance, i.e., the absoluteness, of $L(\mathbb{R})$. Thus, $L(\mathbb{R})^{\mathbb{P}max}$ is (i) effectively complete, i.e. invariant under set-forcing extensions; and (ii) maximal, i.e. satisfies all Π_2-sentences and is thus consistent by set-forcing over ground models (Woodin, ms: 28).

Assume ZFC and that there is a proper class of Woodin cardinals; $A \in \mathbb{P}(\mathbb{R}) \cap L(\mathbb{R})$; ϕ is a Π_2-sentence; and V(G), s.t. $\langle H(\omega_2), \in, I_{NS}, A^G \rangle \models$ 'ϕ': Then, it can be proven that $L(\mathbb{R})^{\mathbb{P}max} \models \langle H(\omega_2), \in, I_{NS}, A^G \rangle \models$ 'ϕ', where 'ϕ' $:= \exists A \in \Gamma^\infty \langle H(\omega_1), \in, A \rangle \models \psi$.

The axiom of determinacy (AD) states that every set of reals, $a \subseteq \omega^\omega$ is determined, where κ is determined if it is decidable.

Woodin's (1999) Axiom (*) can be thus countenanced:

$AD^{L(\mathbb{R})}$ and $L[(\mathbb{P}\omega_1)]$ is a $\mathbb{P}$max-generic extension of $L(\mathbb{R})$,

from which it can be derived that $2^{\aleph_0} = \aleph_2$. Thus, $\neg$CH; and so CH is absolutely decidable.

[8]See Section 5, for further discussion of the duality of knowledge, and its relation to doxastic operators.

By necessitation,

2. K[∃p(p ∧ ¬Kp)] $\iff$ K[ϵp.A(p) ∧ ¬Kϵp.A(p)].

Applying modal axioms, KT, to (1), however,

3. ¬K[ϵp.A(p) ∧ ¬Kϵp.A(p)].

Thus,

4. ¬K∃p(p ∧ ¬Kp).

Leitgeb suggests that (4) be rewritten

5. ⟨K⟩∀p(p → Kp).

Abbreviate (5) by B. By existential introduction and modal axiom K, both

6. B → ∃p[K(p → B) ∨ K(p → ¬B) ∧ p], and

7. ¬B → ∃p[K(p → B) ∨ K(p → ¬B) ∧ p].

Thus,

8. ∃p[K(p → B) ∨ K(p → ¬B) ∧ p].

Abbreviate (8) by C(p). Introducing epsilon notation,

9. [K(ϵp.C(p) → B) ∨ K(ϵp.C(p) → ¬B)] ∧ ϵp.C(p).

By K,

10. [K(ϵp.C(p) → KB) ∨ K(ϵp.C(p) → K¬B)].

From (9) and necessitation, one can further derive

11. Kϵp.C(p).

By (10) and (11),

12. KB ∨ K¬B.

From (5), (12), and K, Leitgeb derives

13. KB.

By, then, the T axiom,

14. ∀p(p → Kp) (291-292).

Rather than accounting for the coextensiveness of epistemic provability and truth, Leitgeb interprets the foregoing result as cause for pessimism with regard to whether the formulas countenanced in epistemic logic and via epsilon terms are genuinely logical truths if true at all (292).

In response to the attending pressure on the status of epistemic logic as concerning truths of logic, one can challenge the derivation, in the above proof, from lines (12) to (13). The inference depends on line (5), i.e., the epistemic possibility of completeness: ⟨K⟩∀p(p → Kp). One can question how, from (4), i.e. the unprovability of the unprovability of a proposition [¬K∃p(p ∧ ¬Kp)], one can derive (5), i.e. that it is epistemically possible that all propositions are informally provable. Assume, however, that line (5) is valid. Then, the validity of the inference from (12) to (13) can be

challenged by the restriction on the quantifier on worlds in the Knowability Principle expressed by (5). The epistemic operator in lines (12) and (13) records, by contrast, the epistemic necessity, rather than the possibility, of the truth of the formulas and subformulas therein. Thus, from (12) either the provability of the provability of propositions or the provability of the unprovability of propositions, one cannot derive (13) the provability of the provability of propositions, because – by (5) – it is only epistemically possible that all true propositions are provable.

11.4   Concluding Remarks

In this paper, I have endeavored to delineate the types of mathematical modality, and to argue that the epistemic interpretation of two-dimensional intensional semantics can be applied in order to explain, in part, the epistemic status of large cardinal axioms and the decidability of Orey sentences. The formal constraints on mathematical conceivability adumbrated in the foregoing can therefore be considered a guide to our possible knowledge of unknown mathematical truth.

Chapter 12

Epistemic Modality, Grothendieck Universes, and Indefinite Extensibility

This essay endeavors to provide a characterization of the notion of definiteness, in order to provide a non-circular definition of the concept of indefinite extensibility. The concept of indefinite extensibility is introduced by Dummett (1963/1978), in the setting of a discussion of the philosophical significance of Gödel's (1931) first incompleteness theorem. Gödel's theorem can be characterized as stating that – relative to a coding defined over the signature of first-order arithmetic, a predicate expressing the property of provability, and a fixed point construction which is non-trivial, such that the formula in which the above predicate figures precludes interpretations such as '0=1' – the formula can be defined as not satisfying the provability predicate. Dummett's concern is with the conditions on our grasp of the concept of natural number, given that the latter figures in a formula whose truth appears to be satisfied despite the unprovability – and thus non-constructivist profile – thereof (186). His conclusion is that the concept of natural number 'exhibits a particular variety of inherent vagueness, namely indefinite extensibility', where a 'concept is indefinitely extensible if, for any definite characterisation of it, there is a natural extension of this characterisation, which yields a more inclusive concept; this extension will be made according to some general principle for generating such extensions, and, typically, the extended characterisation will be formulated by reference to the previous, unextended, characterisation' (195-196). Elaborating on the notion of indefinite extensi-

bility, Dummett (1996: 441) redefines the concept as follows: an 'indefinitely extensible concept is one such that, if we can form a definite conception of a totality all of whose members fall under the concept, we can, by reference to that totality, characterize a larger totality all of whose members fall under it'. Subsequent approaches to the notion have endeavored to provide a more precise elucidation thereof, either by providing an explanation of the property which generalizes to an array of examples in number theory and set theory (cf. Wright and Shapiro, 2006), or by availing of modal notions in order to capture the properties of definiteness and extendability which are constitutive of the concept (cf. Fine, 2006; Linnebo, 2013; Uzquiano, 2015,a). However, the foregoing modal characterizations of indefinite extensibility have similarly been restricted to set-theoretic languages. Furthermore, the modal notions that the approaches avail of are taken to belong to a proprietary type which is irreducible to either the metaphysical or the logical interpretations of the operator.

The aim of this essay is to redress the foregoing, by providing a modal characterization of indefinite extensibility in the setting of category theory, rather than number or set theory. One virtue of the category-theoretic, modal definition of indefinite extensibility is that it provides for a robust account of the epistemological foundations of modal-structuralist approaches to the ontology of mathematics. A second aspect of the philosophical significance of the examination is that it can serve to redress the lacuna noted in the appeal to an irreducible type of mathematical modality, which is argued (i) to be representational, (ii) still to bear on the ontological expansion of domains of sets, and yet (iii) not to range over metaphysical possibilities. By contrast to the latter approach, the category-theoretic characterization of indefinite extensibility is able to identify the functors of coalgebraic non-deterministic automata with elementary embeddings and the modal properties of set-theoretic, Ω-logical consequence.

In Section **2**, I examine the extant approaches to explaining both the property and the understanding-conditions on the concept of indefinite extensibility. In Section **3**, I outline the elements of the category theory of sets and define Grothendieck Universes. In Section **4**, I identify Grothendieck Universes with modal coalgebraic automata, and define the notion of indefinite extensibility in the category-theoretic setting. I argue that the category-theoretic definition of indefinite extensibility, via Grothendieck Universes as modal coalgebraic automata, yields an explanation of the generative property of indefinite extensibility, as well as of the notion of definiteness which

178

figures in the definition. I argue that the generative property of indefinite extensibility can be captured by identifying Kripke functors of colagebras with elementary embeddings. I argue, then, that the notion of definiteness can be captured by the role of Grothendieck Universes-as-modal coalgebraic automata in characterizing the modal profile of Ω-logical consequence, where the latter accounts for the absoluteness of mathematical truths throughout the set-theoretic multiverse. The category-theoretic definition is shown to circumvent the issues faced by rival attempts to define indefinite extensibility via extensional and intensional notions within the setting of set theory. Section **5** provides concluding remarks.

12.1 Indefinite Extensbility in Set Theory: Modal and Extensional Approaches

Characterizations of indefinite extensibility have so far occurred in the language of set theory, and have availed of both extensional and intensional resources. In an attempt to define the notion of definiteness, Wright and Shapiro (op. cit.) argue, for example, that indefinite extensibility may be intuitively characterized as occurring when there is a function which falls under a first-order concept; for a sub-concept of the first-order concept, an application of the function on the sub-concept does not fall within that sub-concept's range; however, a new sub-concept can be formed, and defined as the set-theoretic union of the initial sub-concept and the function applied thereon (266).

Formally, let Π be a higher-order concept of type τ. Let P be a first-order concept falling under Π of type τ. Let f be a function from entities to entities of the same type as P. Finally, let X be a sub-concept of P. P is indefinitely extensible with respect to Π, if and only if:

$\epsilon(P) = f(X)$,

$\epsilon(X) = \neg[f(X)]$, and

$\exists X'[\Pi(X') = (X\cup\{fX\})]$ (op. cit.).

The notion of definiteness is then defined as the limitless preservation of 'Π-hood' by sub-concepts thereof 'under iteration of the relevant operation', f (269).

The foregoing impresses as a necessary condition on the property of indefinite extensibility. Wright and Shapiro note, e.g., that the above formalization

179

generalizes to an array of concepts countenanced in first-order number theory and analysis, including concepts of the finite ordinals (defined by iterations of the successor function); of countable ordinals (defined by countable order-types of well-orderings); of regular cardinals (defined as occurring when the cofinality of a cardinal, κ – comprised of the unions of sets with cardinality less than κ – is identical to κ); of large cardinals (defined by elementary embeddings from the universe of sets into proper subsets thereof, which specify critical points measured by the ordinals); of real numbers (defined as cuts of sets of rational numbers); and of Gödel numbers (defined as natural numbers of a sequence of recursively enumerable truths of arithmetic) (266-267).

As it stands, however, the definition might not be sufficient for the definition of indefinite extensibility, by being laconic about the reasons for which new sub-concepts – comprised as the union of preceding sub-concepts with a target operation defined thereon – are presumed interminably to generate. In response to the above desideratum, concerning the reasons for which indefinite extensibility might be engendered, philosophers have recently appealed to modal properties of the formation of sets. Fine (2006) argues, e.g., that – in order to avoid the Russell property when quantifying over all sets – there are interpretational modalities which induce a reinterpretation of quantifier domains, and serve as a mechanism for tracking the ontological inflation of the hieararchy of sets via, e.g., the power-set operation (2007). Fine (2005) suggests that the interpretational modality at issue might be a species of dynamic modality, which defines modalities as concerning the information entrained by program executions. Reinhardt (1974) and Williamson (2007) argue that modalities are inter-definable with counterfactuals. While Williamson (2016) argues both that imaginative exercises take the form of counterfactual presuppositions and that it is metaphysically possible to decide propositions which are undecidable relative to the current axioms of extensional mathematical languages such as ZF – Reinhardt (op. cit.) argues that large cardinal axioms and undecidable sentences in extensional ZF can similarly be imagined as obtaining via counterfactual presupposition. In an examination of the iterative hierarchy of sets, Parsons (1977/1983) notes that the notion of potential infinity, as anticipated in Book 3, ch. 6 of Aristotle's *Physics*, may be codified in a modal set theory by both a principle which is an instance of the Barcan formula (namely, for predicates P and rigidifying predicates Q, $\forall x(Px \iff Qx) \land \Box\{\forall x(\Box Qx \lor \Box\neg Qx) \land \forall R[\forall x\Box(Qx \to Rx) \to \Box\forall x(Qx \to Rx)]\}$ (fn. 24), as well as a principle for definable set-forming operations (e.g., unions) for Borel sets of reals $\Box(\forall x)\Diamond(\exists y)[y{=}x \cup$

{x}) (528). The modal extension is argued to be a property of the imagination, or intuition, and to apply further to iterations of the successor function in an intensional variant of arithmetic (1979-1980).

Hellman (1990) develops the program intimated in Putnam (1967), and thus argues for an eliminativist, modal approach to mathematical structuralism as applied to second-order plural ZF. The possibilities at issue are taken to be logical – concerning both the consistency of a set of formulas as well as the possible satisfiaction of existential formulas – and he specifies, further, an 'extendability principle', according to which 'every natural model [of ZF] has a proper extension' (421).

Extending Parsons' and Fine's projects, Linnebo (2009, 2013) avails of a second-order, plural modal set theory in order to account for both the notion of potential infinity as well as the notion of definiteness. Similarly to Parsons' use of the Barcan formula (i.e., $\forall \Box \phi \rightarrow \Box \forall \phi$), Linnebo's principle for the foregoing is as follows: $\forall u(u \prec xx \rightarrow \Box \phi) \rightarrow \Box \forall u(u \prec \rightarrow \phi)$ (2013: 211). He argues, further, that the logic for the modal operator is S4.2, i.e. K $[\Box(\phi \rightarrow \psi) \rightarrow (\Box \phi \rightarrow \Box \psi)]$, T $(\Box \phi \rightarrow \phi)$, 4 $((\Box \phi \rightarrow \Box \Box \phi)$, and G $(\Diamond \Box \phi \rightarrow \Box \Diamond \phi)$. Studd (2013) examines the notion of indefinite extensibility by availing of a bimodal temporal logic. Uzquiano's (2015) approach to defining the concept of indefinite extensibility argues that the height of the cumulative hierarchy is in fact fixed, and that indefinite extensibility can similarly be captured via the use of modal operators in second-order plural modal set theory. The modalities are taken to concern the possible reinterpretations of the intensions of the non-logical vocabulary – e.g., the set-membership relation – which figures in the augmentation of the theory with new axioms and the subsequent climb up the fixed hierarchy of sets (cf. Gödel, 1947/1964).

In chapter **11**, I proffered a novel epistemology of mathematics, based on an application of the epistemic interpretation of two-dimensional intensional semantics in set-theoretic languages to the values of large cardinal axioms and undecidable sentences. Modulo logical constraints such as consistency and generic absoluteness in the extensions of ground models of the set-theoretic multiverse, the epistemic possibility that an undecidable proposition receives a value may serve, then, as a guide to the metaphysical possibility thereof. Finally, in chapters **10** and **12**, I argued that the modal profile of the consequence relation, in the Ω-logic defined in Boolean-valued models of set-theory, can be captured by coalgebraic modal automata, and provides a necessary condition on the formal grasp of the concept of 'set'.

The foregoing accounts of the metaphysics and epistemology of indefinite

extensibility are each defined in the languages of number and set theory. In the following section, I examine the nature of indefinite extensibility in the setting of category theory, instead. One aspect of the philosophical significance of the examination is that it can serve to provide an analysis of the mathematical modality at issue, by availing only of model-theoretic resources. By contrast to Hellman's approach, which takes the mathematical modality at issue to be logical (cf. Field, 1989: 37; Rayo, 2013), and Fine's (op. cit.) approach, which takes the mathematical modality to be dynamic, I argue in the following sections that the mathematical modality can be captured by the functors of coalgebraic modal automata, where the latter are identifiable with Grothendieck Universes.

12.2   Grothendieck Universes

We work within a two-sorted language in which the Eilenberg - Mac Lane Axioms of category theory are specified. Types are labeled A,B,C for objects and x,y,z for arrows. The relevant operators are the domain operator, `Dom`, which takes arrows to objects; the codomain operator, `Cod`, which operates similarly, and the identity operator, 1_α, which takes objects to arrows. Finally, a composition relation, C(x,y; z), is defined on arrows, where the open formula reads z is the composite of x and y (McLarty, 2008: 13). The Eilenberg - Mac Lane axioms can then be defined as follows:

- Axioms of Domain and Codomain:

  $\forall$f,g,h, if C(f,g,h), then `Dom`f = `Dom`h and `Cod`f = `Dom`g and `Cod`g = `Cod`h

- Axioms of Existence and Uniqueness of Composites:

  $\forall$f,g, if `Cod`f = `Dom`g, then $\exists$!h, s.t. C(f,g; h)

- Axioms for Identity Arrows:

  $\forall$A, `Dom`1_A = `Cod`1_A = A

  $\forall$f, C($1_{(\mathtt{Dom}f)}$,f; f)

  $\forall$f, C(f,$1_{(\mathtt{Dom}f)}$; f)

- Axiom of Associativity of Composition:

  $\forall$f,g,h,i,j,k, if C(f,g; i) and C(g,h; j) and C(f,j; k), then C(i,h; k) (op. cit.).

182

Categorical Set Theory is defined by augmenting the Eilenberg - Mac Lane axioms with the axioms of Lawvere's Elementary Theory for the Category of Sets (ETCS) (op. cit.; Lawvere, 2005). Following McLarty, we define the singleton of a set as one for which 'every set has exactly one function to it' (op. cit.: 25). An element of a set A, x∈A, is a function x: $1 \to A$ (op. cit.). Composition occurs if and only if, for two arrows, f,g, and object x, (gf)(x) = g(f(x)) (26) The axioms are then defined as follows (op. cit.). Finally, an equalizer e: $E \to A$ for a pair of functions f,g: $A \to B$ is defined as 'a universal solution to the equaltion fe = ge' (29):

- Every pair of sets, A,B, has a product:

  $\forall$T,f,g, with f: $T \to A$, g: $T \to B$, $\exists!\langle f,g \rangle$: $T \to A \times B$

- Every parallel pair of functions, f,g: $A \to B$, has an equalizer:

  $\forall$T,h, with fh = gh, $\exists!$ u: $T \to E$

- There is a function set from each set A to each set B:

  $\forall$C and g: $C \times A \to B$, $\exists!$g': $C \to B^A$

- There is a truth value *true*: $1 \to 2$:

  $\forall$A and monic S $\rightarrowtail$ A, $\exists! \chi_i$, such that S is an equalizer

- There is a natural number triple, $\mathbb{N}$, 0, s:

  $\forall$T and x: $1 \to T$ and f: $T \to T$, $\exists!$ u: $\ltimes \to T$

- Extensionality

  $\forall$f $\neq$ g: $A \to B$, $\exists$x: $1 \to A$, with f(x) $\neq$ g(x)

- Non-triviality

  $\exists false$: $1 \to 2$, s.t. $false \neq true$

- Choice

  $\forall$ onto functions f: $A \to B$, $\exists$h: $B \to A$, s.t. fh = 1_A.

From the axioms of ETCS, only a version of the ZF separation axiom with bounded quantifiers can be recovered (37). The axiom of separation states that $\exists$x$\forall$u[u∈x $\iff$ u∈a $\wedge$ ϕ(u)]. In order to redress the restriction to bounded quantifiers, we work within a stronger category theory for sets,

183

i.e. the 'category of categories as foundation' (CCAF). The axioms of the CCAF build upon those of both ETCS and Eilenberg - Mac Lane category theory, by augmenting them with the following (53):

- Every category C has a unique functor, $C \to 1$

- The category 2 has exactly two functors from 1 and three to itself

- Let a pushout be defined such that if f: $A \to C$ and g: $B \to C$, then a: $C \to A$ and b: $C \to B$ (Pettigrew, ms: 19). The category 3 is a pushout, and there is a functor γ: $2 \to 3$, with $\gamma_0 = \alpha_0$ and $\gamma_1 = \beta_1$

- Arrow Extensionality

  $\forall F,G$: $A \to B$, if $F \neq G$ then $\exists f$: $2 \to A$ with $Ff \neq Gf$.

A Grothendieck Universe may finally be defined as a set, U, which satisfies the axioms of ZF set theory without choice, yet as augmented by at least strongly inaccessible large cardinals. The axioms of ZF are:

- Empty set:

  $\exists x \forall u(u \notin x)$

- Extensionality:

  $x = y \iff \forall u(u \in x \iff u \in y)$

- Pairing:

  $\exists x \forall u(u \in x \iff u = a \lor u = b)$

- Union:

  $\exists x \forall u[u \in x \iff \exists v(u \in v \land v \in a)]$

- Separation:

  $\exists x \forall u[u \in x \iff u \in a \land \phi(u)]$

- Power Set:

  $\exists x \forall u(u \in x \iff u \subseteq a)$

- Infinity:

  $\exists x \varnothing \in x \land \forall u(u \in x \to \{u\} \in x)$

184

- Replacement:

  $\forall u \exists! v \psi(u,v) \rightarrow \forall x \exists y (\forall u \in x)(\exists v \in y)\psi(u,v)$.

Large cardinal axioms are defined by elementary embeddings.[1] Elementary embeddings can be defined thus. For models A,B, and conditions ϕ, j: A $\rightarrow$ B, $\phi\langle a_1, \ldots, a_n \rangle$ in A if and only if $\phi\langle j(a_1), \ldots, j(a_n) \rangle$ in B (Kanamori, 2012: 363). A cardinal κ is regular if the cofinality of κ – comprised of the unions of sets with cardinality less than κ – is identical to κ (op. cit.: 360). Uncountable regular limit cardinals are weakly inaccessible (op. cit.). A strongly inaccessible cardinal is regular and has a strong limit, such that if $\lambda < \kappa$, then $2^\lambda < \kappa$ (op. cit.).

By augmenting languages of the theory of CCAF with Grothendiek Universes, U, CCAF proves thereby that:

CCAF $\vdash \forall n \in \mathbb{N}, \exists\{\aleph_0, \aleph_1, \ldots, \aleph_n\}$, in the category of Sets, U-Set (37-38).

12.3 Modal Coalgebraic Automata and Indefinite Extensibility

This section examines, finally, the reasons for which Grothendieck Universes provide a more theoretically adequate model of the understanding-conditions for mathematical concepts than do competing approaches such as the Neo-Fregean epistemology of mathematics. According, e.g., to the Neo-Fregean program, concepts of number in arithmetic and analysis are definable via implicit definitions which take the form of abstraction principles. Abstraction principles specify biconditionals in which – on the left-hand side of the formula – an identity is taken to hold between numerical term-forming operators from second-order entities to abstract objects, and – on the right-hand side of the formula – an equivalence relation on lower-order entities is assumed to hold.

In the case of cardinal numbers, the relevant abstraction principle is referred to as Hume's principle, and states that, for all x and y, the number of the x's is identical to the number of the y's if and only if the x's and the y's can be put into a one-to-one correspondence, i.e., there is a bijection from the x's onto the y's. Abstraction principles for the concepts of other numbers have further been specified. Thus, e.g., Shapiro (2000: 337-340) specifies

---

[1]Cf. Koellner and Woodin (2010); Woodin (2010).

an abstraction principle for real numbers, which proceeds along the method of Dedekind's definition of the reals (cf. Wright, 2007: 172). According to the latter method, one proceeds by specifying an abstraction principle which avails of the natural numbers, in order to define pairs of finite cardinals: '$\forall x,y,z,w[\langle x,y \rangle = \langle z,w \rangle \iff x = z \land y = w]$. A second abstraction principle is defined which takes the differences of the foregoing pairs of cardinals, identifying the differences with integers: $[\text{Diff}(\langle x,y \rangle) = \text{Diff}(\langle z,w \rangle) \iff x + w = y + z]$. One specifies, then, a principle for quotients of the integers, identifying them subsequently with the rational numbers: $[Q\langle m,n \rangle = Q\langle p,q \rangle \iff n = 0 \land q = 0 \lor n \neq 0 \land q \neq 0 \land m \text{ x } q = n \text{ x } p]$. Finally, one specifies sets of rational numbers, i.e. the Dedekind cuts thereof, and identifies them with the reals: $\forall F,G[\text{Cut}(F) = \text{Cut}(G) \iff \forall r(F \leq r \iff G \leq r)]$.

The abstractionist program faces several challenges, including whether conditions can be delineated for the abstraction principles, in order for the principles to avoid entraining inconsistency[2]; whether unions of abstraction principles can avoid the problem of generating more abstracts than concepts (Fine, 2002); and whether abstraction principles can be specified for mathematical entities in branches of mathematics beyond first and second-order arithmetic (cf. Boolos, 1997; Hale, 2000; Shapiro, op. cit.; and Wright, 2000). I will argue that the last issue – i.e., being able to countenance definitions for the entities and structures in branches of mathematics beyond first and second-order arithmetic – is a crucial desideratum, the satisfaction of which remains elusive for the Neo-Fregean program while yet being satisfiable and thus adducing in favor of the modal platonist approach that is outlined in what follows.

One issue for the attempt, along abstractionist lines, to provide an implicit definition for the concept of set is that doing so with an unrestricted comprehension principle yields a principle identical to Frege's (1893/2013) Basic Law V; and thus – in virtue of Russell's paradox – entrains inconsistency. However, two alternative formulas can be defined, in order to provide a suitable restriction to the inconsistent abstraction principle. The first, conditional principle states that $\forall F,G[[\text{Good}(F) \lor \text{Good}(G)] \rightarrow [\{x|Fx\} = \{Gx\} \iff \forall x(Fx \iff Gx)]]$. The second principle is an unconditional version of the foregoing, and states that $\forall F,G[\{x|Fx\} = \{Gx\} \iff [\text{Good}(F) \lor$

[2]Cf. Hodes (1984,a); Hazen (1985); Boolos (1990); Heck (1992); Fine (2002); Weir (2003); Cook and Ebert (2005); Linnebo and Uzquiano (2009); Linnebo (2010); and Walsh (forthcoming).

Good(G) → ∀x(Fx ⟺ Gx)]]. Following von Neumann's (1925/1967: 401-402) suggestion that Russell's paradox can be avoided with a restriction of the set comprehension principle to one which satisfies a constraint on the limitation of its size, Boolos (1997) suggests that the 'Good' predicate in the above principles is intensionally isomorphic to the notion of smallness in set size, and refers to the principle as New V. However, New V is insufficient for deriving all of the axioms of ZF set theory, precluding, in particular, both the axioms of infinity and the power-set axiom (cf. Wright and Hale, 2005: 193). Further, there are other branches of number theory for which it is unclear whether acceptable abstraction principles can be specified. Wiles' proof of Fermat's Last Theorem (i.e., that, save for when one of the variables is 0, the Diophantine equation, $x^n = y^n = z^n$, has no solutions when n $>$2; cf. Hardy and Wright, 1979: 190) relies, e.g., on both invariants and Grothendieck universes in cohomological number theory (cf. McLarty, 2009: 4).

The foregoing issues with regard to the definability of abstracta in number theory, algebraic geometry (McLarty, op. cit.: 6-8), set theory, et al., can be circumvented in the category-theoretic setting; and in particular by colagebras. In the remainder of this section, I endeavor to demonstrate how modal coalgebraic automata are able to countenance two, fundamental mathematical concepts. The first is the target concept in this essay, namely indefinite extensibility. The second concerns the epistemic and modal properties of the concept of logical consequence, in the Ω-logic in axiomatic set theory.

A labeled transition system is a tuple, LTS, comprised of a set of worlds, M; a valuation, V, from M to its powerset, P(M); and a family of accessibility relations, R. So LTS = ⟨M,V,R⟩ (cf. Venema, 2012: 7). A Kripke coalgebra combines V and R into a Kripke functor, σ_R; i.e. the set of binary morphisms from M to P(M) (op. cit.: 7-8). Thus, for an s∈M, σ(s) := [σ_V(s), σ_R(s)] (op. cit.). Satisfaction for the system is defined inductively as follows: For a formula ϕ defined at a state, s, in M,

$[\![\phi]\!]^M$ = V(s) [3]
$[\![\neg\phi]\!]^M$ = S − V(s)
$[\![\bot]\!]^M$ = ∅
$[\![T]\!]^M$ = M
$[\![\phi \vee \psi]\!]^M = [\![\phi]\!]^M \cup [\![\psi]\!]^M$
$[\![\phi \wedge \psi]\!]^M = [\![\phi]\!]^M \cap [\![\psi]\!]^M$
$[\![\diamond_s\phi]\!]^M = \langle R_s \rangle [\![\phi]\!]^M$

---

[3]Equivalently, M,s ⊩ ϕ if s∈V(ϕ) (9).

187

$[\![\Box_s\phi]\!]^M = [\mathrm{R}_s][\![\phi]\!]^M$, with

$\langle\mathrm{R}_s\rangle(\phi) := \{\mathrm{s}'\in\mathrm{S} \mid \mathrm{R}_s[\mathrm{s}'] \cap \phi \neq \emptyset\}$ and

$[\mathrm{R}_s](\phi) := \{\mathrm{s}'\in\mathrm{S} \mid \mathrm{R}_s[\mathrm{s}'] \subseteq \phi\}$ (9).

Kripke Coalgebras can be identified with Grothendieck Universes. In CCAF, the elementary embeddings which are jointly necessary and sufficient for positing the existence of large cardinals can further be identified with the functors, i.e. transition functions, in modal coalgebraic automata. Finally, Kripke coalgebras are the dual representations of Boolean-valued models of the Ω-logic of set theory (cf. Venema, 2007). When identified with Grothendieck Universes, modal coalgebraic automata are able, then, to countenance the consitutive conditions of indefinite extensibility. Modal coalgebraic automata are capable, e.g., of defining both the generative property of indefinite extensibility, as well as the notion of definiteness which figures therein. Further, the category-theoretic definition of indefinite extensibility is arguably preferable to those advanced in the set-theoretic setting, because modal coalgebraic automata can account for both the modal profile and the epistemic tractability of Ω-logical consequence.

The *generative* property of indefinite extensibility is captured by the foregoing Kripke functor, σ_R – i.e., the morphism mapping a Boolean-valued model to the powerset thereof – and which we have identified with elementary embeddings, j: A $\to$ B, $\phi\langle\mathrm{a}_1, \ldots, \mathrm{a}_n\rangle$ in A if and only if $\phi\langle\mathrm{j}(\mathrm{a}_1), \ldots, \mathrm{j}(\mathrm{a}_n)\rangle$ in B.

The notion of *definiteness* is captured by the role of modal coalgebraic automata in characterizing the modal profile of Ω-logical validity. Ω-logical validity can be defined as follows:

For T$\cup\{\phi\}\subseteq Sent$,

T $\models_\Omega \phi$, if for all ordinals a and countable Boolean algebras $\mathbb{B}$, if $\mathrm{V}^{\mathbb{B}}{}_a \models$ T, then $\mathrm{V}^{\mathbb{B}}{}_a \models \phi$ (Bagaria et al., 2006). The Ω-Conjecture states that V $\models_\Omega \phi$ iff $\mathrm{V}^{\mathbb{B}} \models_\Omega \phi$ (Woodin, ms). Thus, Ω-logical validity is invariant in all set-forcing extensions of ground models in the set-theoretic multiverse.

The invariance property of Ω-logical consequence can then be characterized by modal coalgebraic automata. Mathematical truths are thus said to be definite in virtue of holding of necessity, as recorded by the functors of the modal colagebraic automata which are dually isomorphic to the Boolean-valued algebraic models for the Ω-logic of set theory.

Thus, whereas the Neo-Fregean approach to comprehension for the concept of set relies on an unprincipled restriction of the size of the universe in order to avoid inconsistency, and one according to which the axioms of ZF

still cannot all be recovered, modal coalgebraic automata provide a natural means for defining the minimal conditions necessary for formal grasp of the concept set. The category-theoretic definition of indefinite extensibility is sufficient for uniquely capturing both the generative property as well as the notion of definiteness which are constitutive of the concept. Finally, a further point adducing in favor of the category-theoretic definition of indefinite extensibility is that it requires no appeal to a notion of mathematical modality which problematically endeavors to capture both the epistemic property of possible interpretations of quantifiers, as well as the metaphysical property of set-theoretic ontological expansion.

12.4   Concluding Remarks

In this essay, I outlined a number of approaches to defining the notion of indefinite extensibility, each of which restricts the scope of their characterization to set-theoretic languages. I endeavored, then, to define indefinite extensibility in the setting of category-theoretic languages, and examined the benefits accruing to the approach, by contrast to the extensional and modal approaches pursued in ZF.

The extensional definition of indefinite extensibility in ZF was shown to be insufficient for characterizing the generative property in virtue of which number-theoretic concepts are indefinitely extensible. The generative property of indefinite extensibility in the category-theoretic setting was argued, by contrast, to be identifiable with the Kripke functors of modal coalgebraic automata, where the automata model Grothendieck Universes, and Kripke functors are further identifiable with the elementary embeddings by which large cardinal axioms can be specified. The modal definitions of indefinite extensibility in ZF were argued to be independently problematic, in virtue of endeavoring simultaneously to account for the epistemic properties of indefinite extensibility – e.g., possible reinterpretations of quantifier domains and mathematical vocabulary – as well as the metaphysical properties of indefinite extensibility – i.e., the ontological expansion of the target domains. By contrast, the Kripke functors definable in Grothendieck universes-as-modal coalgebraic automata were argued to circumvent the foregoing conflation, accounting just for the ontological expansion effected by elementary embeddings in the category of sets.

Finally, against the Neo-Fregean approach to defining concepts of num-

ber, and the limits thereof in the attempt to define concepts of mathematical objects in other branches of mathematics beyond arithmetic, I demonstrated how – by characterizing the modal profile of Ω-logical validity and thus the generic invariance and absoluteness of mathematical truths concerning large cardinals throughout the set-theoretic multiverse – modal coalgebraic automata are capable of capturing the notion of definiteness within the concept of indefinite extensibility.

Chapter 13

A Modal Logic for Gödelian Intuition

> 'The incompleteness results do not rule out the possibility that there is a theorem-proving computer which is in fact equivalent to mathematical intuition' – Gödel, quoted in Wang (1986: 186).[1]

13.1   Introduction

Gödel's writings on the epistemology of mathematics number fewer than twenty pages. He avails, in his remarks, of a notion of non-sensory intuition – alternatively, 'consciousness', or 'phenomenology' (cf. Gödel, 1961: 383) – as a fundamental, epistemic conduit into mathematical truths.[2] According to Gödel, the defining properties of mathematical intuition include (i) that it either is, or is analogous to, a type of perception (1951: 323; 1953,V: 359;

---

[1] Note however that, in the next subsequent sentence, Gödel records scepticism about the foregoing. He remarks: 'But they imply that, in such a – highly unlikely for other reasons – case, either we do not know the exact specification of the computer or we do not know that it works correctly' [Gödel, quoted in Wang (op. cit.)].

[2] Another topic that Gödel suggests as being of epistemological significance is the notion of 'formalism freeness', according to which the concepts of computability, demonstrability (i.e., absolute provability), and ordinal definability can be specified independently of a background formal language (cf. Gödel 1946, and Kennedy 2013 for further discussion). Kennedy notes however that, in his characterizations of demonstrability and definability, Gödel assumes ZFC as his metatheory (op. cit.: 383). Further examination of the foregoing is beyond the scope of the present essay.

1964: 268); (ii) that it enables subjects to alight upon new axioms which are possibly true (1953,III: 353,fn.43; 1953,V: 361; 1961: 383, 385; 1964: 268); (iii) that it is associated with modal properties, such as provability and necessity (1933: 301; 1964: 261); and (iv) that the non-sensory intuition of abstracta such as concepts entrains greater conceptual 'clarification' (1953,III: 353,fn.43; 1961: 383). Such intuitions are purported to be both *of* abstracta and formulas, as well as to the effect *that* the formulas are true.[3]

In this paper, I aim to outline the logical foundations for rational intuition, by examining the nature of property (iii). The primary objection to Gödel's approach to mathematical knowledge is that the very idea of rational intuition is insufficiently constrained.[4] Subsequent research has thus endeavored to expand upon the notion, and to elaborate on intuition's roles. Chudnoff (2013) suggests, e.g., that intuitions are non-sensory experiences which represent non-sensory entities, and that the justificatory role of intuition is that it enables subjects to be aware of the truth-makers for propositions (p. 3; ch. 7). He argues, further, that intuitions both provide evidence for beliefs as well as serve to guide actions (145).[5] Bengson (2015: 718-723) suggests that rational intuition can be identified with the 'presentational', i.e., phenomenal, properties of representational mental states – namely, cognitions – where the phenomenal properties at issue are similarly non-sensory; are not the product of a subject's mental acts, and so are 'non-voluntary'; are qualitatively gradational; and they both 'dispose or incline assent to their contents' and further 'rationalize' assent thereof.[6]

---

[3]The distinction between 'intuition-of' and 'intuition-that' is explicitly delineated in Parsons (1980: 145-146), and will be further discussed in Section **2**.

[4]See, e.g., Hale and Wright (2002). Wright (2004) provides a vivid articulation of the issue: 'A major — but not the only – problem is that, venerable as the tradition of postulating intuitive knowledge of first principles may be, no-one working within it has succeeded at producing even a moderately plausible account of how the claimed faculty of rational intuition is supposed to work — how exactly it might be constituted so as to be reliably responsive to basic logical validity as, under normal circumstances, vision, say, is reliably responsive to the configuration of middle-sized objects in the nearby environment of a normal human perceiver' (op. cit.: 158).

[5]A similar proposal concerning the justificatory import of cognitive phenomenology – i.e., the properties of consciousness unique to non-sensory mental states such as belief – can be found in Smithies (2013a,b). Smithies prescinds, however, from generalizing his approach to the epistemology of mathematics.

[6]Compare Kriegel (2015: 68), who stipulates that 'making a judgment that p involves a feeling of involuntariness' and 'making a judgment always involves the feeling of mobilizing a concept'.

Rather than target objections to the foregoing essays, the present discussion aims to rebut the primary objection to mathematical intuition alluded to above, by providing a logic for its defining properties. The significance of the proposal is thus that it will make the notion of intuition formally tractable, and might thus serve to redress the contention that the notion is mysterious and ad hoc.

In his (1933) and (1964), Gödel suggests that intuition has a constitutively modal profile. Constructive intuitionistic logic is shown to be translatable into the modal logic, S4, with the rule of necessitation, while the modal operator is interpreted as concerning provability.[7] Mathematical intuition of set-theoretic axioms is, further, purported both to entrain 'intrinsic' justification, and to illuminate the 'intrinsic necessity' thereof.[8] Following Gödel's line of thought, I aim, in this paper, to provide a modal logic for the notion of 'intuition-that'.[9]

If rational intuition is identical with cognitive phenomenal properties of representational states such as beliefs and judgments, then – via correspondence results between modal logic and first-order logic [cf. van Benthem (1983; 1984/2003)] – a precise translation can be provided between the notion of 'intuition-of', i.e., the cognitive phenomenal properties of thoughts whose contents can concern the axioms of mathematical languages, and the

---

[7]For further discussion both of provability logic and of intuitionistic systems of modal logic, see Löb (1955); Smiley (1963); Kripke (1965); and Boolos (1993). Löb's provability formula was formulated in response to Henkin's (1952) problem concerning whether a sentence which ascribes the property of being provable to itself is provable. (Cf. Halbach and Visser, 2014, for further discussion.) For an anticipation of the provability formula, see Wittgenstein (1933-1937/2005: 378), where Wittgenstein writes: 'If we prove that a problem can be solved, the concept 'solution' must somehow occur in the proof. (There must be something in the mechanism of the proof that corresponds to this concept.) But the concept mustn't be represented by an external description; it must really be demonstrated. / The proof of the provability of a proposition is the proof of the proposition itself' (op. cit.). Wittgenstein distinguishes the foregoing type of proof with 'proofs of relevance' which are akin to the mathematical, rather than empirical, propositions, discussed in Wittgenstein (2001: IV, 4-13, 30-31).

[8]Gödel (1964) does not define intrinsic justifications, although he does contrast the latter with the notion of extrinsic justifications, for which he provides a few defining remarks. Extrinsic justifications are associated, for example, with both the evidential probability of propositions, and the 'fruitful' consequences of a mathematical theory subsequent to adopting new axioms. See Gödel (op. cit.: 269).

[9]Cf. Parsons (1979-1980; 1983: p. 25, chs.10-11; 2008: 176).

modal operators regimenting the notion of 'intuition-that'.[10] I argue, then, that intuition-that can further be shown to entrain property (iv), i.e. conceptual elucidation, by way of figuring as an interpretational modality which induces the reinterpretation of domains of quantification (cf. Fine, 2005; 2006). Fine (op. cit.) has suggested that the interpretational modality is imperatival, and that the deontic aspects of the modality might best be captured by a dynamic logic (p.c.). Following Fine's suggestion, I argue that intuition-that can thus be understood to be a species of fixed point dynamic provability logic, which is equivalent to the bisimulation-invariant fragment of monadic second-order logic (cf. Janin and Walukiewicz, 1996; Venema, 2014, ms). Modalized rational intuition is therefore expressively equivalent to – and can crucially serve as a guide to the interpretation of – the entities, such as mathematical concepts, that are formalizable in monadic first- and second-order formal languages.

In Section **2**, I elucidate the properties of rational intuition, by examining arguments and evidence adducing in favor of the existence of cognitive phenomenal consciousness. In Section **3**, I countenance and motivate a multi-modal logic, which augments the provability logic, GL, with fixed point dynamic logic, i.e. the modal μ-calculus. I argue that the dynamic properties of modalized rational intuition provide a precise means of accounting for the manner by which intuition can yield the reinterpretation of quantifier domains and mathematical vocabulary; and thus explain the role of rational intuition in entraining conceptual elucidation. In Section **4**, I examine remaining objections to the viability of rational intuition and provide concluding remarks.

13.2 Rational Intuition as Cognitive Phenomenology

A property of a mental state is phenomenal only if it is the property of being aware of the state. If the mental state at issue is sensory, then sensory phenomenal properties will be properties of being aware of one's perceptions, where the perceptions at issue will be unique to the subject's sensory modal-

---

[10]This provides a precise answer to the target inquiry advanced by Parsons (1993: 233).

ities of vision, audition, etc.[11] Let cognitive phenomenal consciousness refer to the properties of being aware of the non-sensory representational mental states toward which subjects can bear attitudes. Such states can be identified with, e.g., formulas, ϕ, in a Language of Thought,[12] the syntax for which mirrors that of natural language sentences, and which can fall within the scope of various operators such as fully or partially believing that (x, x a propositional variable/ϕ), knowing that (x/ϕ), judging that (x/ϕ), asserting that (x/ϕ), questioning whether (x/ϕ) has a particular semantic value, et al.

Pitt (2004: 8) provides the following argument for the existence of cognitive phenomenal properties, which – for the purposes of this essay – I will assume to be sound:

(K1) 'It is possible immediately to identify one's occurrent conscious thoughts (equivalently (see below): one can know by acquaintance which thought a particular occurrent conscious thought is); but

(K2) It would not be possible immediately to identify one's conscious thoughts unless each type of conscious thought had a proprietary, distinctive, individuative phenomenology; so

(P) Each type of conscious thought – each state of consciously thinking that p, for all thinkable contents p – has a proprietary, distinctive, individuative phenomenology'.

In his examination of the conditions on measuring partial beliefs, i.e., subjective probability, Ramsey (1926) records scepticism about whether subjects are aware in a non-sensory way of all of their (partial) beliefs.[13] For the pur-

---

[11]For issues concerning the taxonomy of the sense modalities, see Macpherson (2011). Bottom-up/exogenous, spatial-based, property-based, and diffuse/focal attention arguably comprise a necessary condition on the instantiation of phenomenal properties. The condition is witnessed by the phenomenon referred to as the attentional blink. The attentional blink holds if and only if shifting attentional allocation from one stimulus to another induces a lack of awareness of the first stimulus to which attention was previously distributed. See chapter **6** and the essays in Mole et al. (eds.) (2010), for further discussion.

[12]Cf. Fodor (1975).

[13]See Ramsey (op. cit.: 169): 'Suppose that the degree of a belief is something perceptible by its owner; for instance that beliefs differ in the intensity of a feeling by which they are accompanied, which might be called a belief-feeling or feeling of conviction, and that by the degree of belief we mean the intensity of this feeling. This view would be very inconvenient, for it is not easy to ascribe numbers to the intensities of feelings; but apart from this it seems to me observably false, for the beliefs which we hold most strongly are often accompanied by practically no feeling at all; no one feels strongly about things he takes for granted'. Contrast Koopman (1940: 271), who argues that intuition can serve as a guide to veridical judgments concerning the comparison of probability measures [cf.

poses of this note, it is sufficient that at least some cognitive states are states of which subjects can be aware – where, again, the properties of awareness at issue are non-sensory and purported to be unique to distinct cognitive attitudes, such that being aware of one's belief that ϕ will qualitatively differ from one's awareness of one's interrogative state concerning whether ϕ is true.

The evidence for the claim that at least some cognitive states are associated with a unique set of non-sensory properties of awareness has proceeded via introspective reports.[14] In the latter, subjects verbally report upon the valence of their awareness of their states, where their reports are assumed to be reliable.[15] One phenomenon of awareness of one's thoughts might, e.g., be that of inner speech.[16] It is an open issue whether inner speech has a sensory basis (cf. Prinz, 2012). If not, however, and assuming that introspective report is a reliable method for discerning whether subjects are aware of their states – irrespective of their being able to ascertain a precise value thereof – then there might be properties of awareness that are unique to one's thoughts and cognitive propositional attitudes.[17]

de Finetti (1937/1964: ch.1) on the primacy of comparative judgments of probability], as well as to the axioms – or laws of consistency – of a probability theory.

[14]The results of the method of introspection are availed of by Pitt (op. cit.), and discussed in the essays in Bayne and Montague (eds.) (2011). For an excellent survey of the experimental paradigms endeavoring to corroborate that intuition can be a source of evidence, see Nagel (2007; 2013a).

[15]See, however, Schwitzgebel (2011) for an examination of a series of case studies evincing that introspective report is unreliable as a method for measuring consciousness.

[16]Cf. Carruthers (1996). Assuming the reliability of introspection, Machery (2005) argues that awareness of one's thoughts is nevertheless insufficient for ascertaining the syntactic structure thereof.

[17]Nagel (2013) examines an approach to intuitions which construes the latter as a type of cognition, rather than as a phenomenal property of judgments. She distinguishes, e.g., between intuition and reflection, on the basis of experimental results which corroborate that there are distinct types of cognitive processing (op. cit.: 226-228). Intuitive and reflective cognitive processing are argued to interact differently with the phenomenal information comprising subjects' working memory stores. Nagel notes that – by contrast to intuitive cognition – reflective cognition 'requires the sequential use of a progression of conscious contents to generate an attitude, as in deliberation' (231).

13.3 Modalized Rational Intuition and Conceptual Elucidation

In this section, I will outline the logic for Gödelian intuition. The motivation for providing a logic for rational intuition will perhaps be familiar from treatments of the property of knowledge in formal epistemology. The analogy between rational intuition and the property of knowledge is striking: Just as knowledge has been argued to be a mental state (Williamson, 2001; Nagel, 2013b); to be propositional (Stanley and Williamson, 2001); to be factive; and to possess modal properties (Hintikka, 1962; Nozick, 1981; Fagin et al., 1995; Meyer and van der Hoek, 1995), so rational intuition can be argued to be a property of mental states; to be propositional, as recorded by the notion of intuition-that; and to possess modal properties amenable to rigorous treatment in systems of modal logic.

I should like to suggest that the modal logic of Gödelian intuition is the bimodal logic combining GL – which is comprised of axioms K, 4, GL, and the rule of necessitation – with the modal μ-calculus.

Let M be a model over the Kripke frame, $\langle$W, R$\rangle$; so, M = $\langle$W, R, V$\rangle$. W is a non-empty set of possible worlds. R is a binary relation on W. V is a function assigning proposition letters, ϕ, to subsets of W.

$\langle$M,w$\rangle \Vdash \phi$ if and only if w$\in$V(ϕ).

$\langle$M,w$\rangle \Vdash \neg\phi$ iff it is not the case that $\langle$M,w$\rangle \Vdash \phi$

$\langle$M,w$\rangle \Vdash \phi \wedge \psi$ iff $\langle$M,w$\rangle \Vdash \phi$ and $\langle$M,w$\rangle \Vdash \psi$

$\langle$M,w$\rangle \Vdash \phi \vee \psi$ iff $\langle$M,w$\rangle \Vdash \phi$ or $\langle$M,w$\rangle \Vdash \psi$

$\langle$M,w$\rangle \Vdash \phi \rightarrow \psi$ iff, if $\langle$M,w$\rangle \Vdash \phi$, then $\langle$M,w$\rangle \Vdash \psi$

$\langle$M,w$\rangle \Vdash \phi \iff \psi$ iff [$\langle$M,w$\rangle \Vdash \phi$ iff $\langle$M,w$\rangle \Vdash \psi$]

$\langle$M,w$\rangle \Vdash \diamond\phi$ iff $\exists$w'[R(w,w') and $\langle$M,w'$\rangle \Vdash \phi$]

$\langle$M,w$\rangle \Vdash \Box\phi$ iff $\forall$w'[If R(w,w'), then $\langle$M,w'$\rangle \Vdash \phi$].

K states that $\Box(\phi\rightarrow\psi) \rightarrow (\Box\phi\rightarrow\Box\psi)$; i.e., if one has an intuition that ϕ entails ψ, then if one has the intuition that ϕ then one has the intuition that ψ. GL states that $\Box[(\Box\phi \rightarrow \phi) \rightarrow \Box\phi]$; i.e., if one has the intuition that the intuition that ϕ entails that ϕ is true, then one has the intuition that ϕ. 4 states that $\Box\phi \rightarrow \Box\Box\phi$; i.e., if one has the intuition that ϕ, then one intuits that one has the intuition that ϕ. Necessitation states that $\vdash\phi \rightarrow \vdash\Box\phi$. Because intuition-that is non-factive, we eschew in our modal system of axiom T, which states that $\Box\phi \rightarrow \phi$; i.e., one has the intuition that ϕ only if ϕ is the case [cf. BonJour (1998: 4.4); Parsons (2008: 141)].

In order to account for the role of rational intuition in entraining conceptual elucidation (cf., Gödel, 1961: 383), I propose to follow Fine (2006) and Uzquiano (2015,a) in suggesting that there are interpretational modalities associated with the possibility of reinterpreting both domains of quantification (Fine, op. cit.) and the non-logical vocabulary of mathematical languages, such as the membership relation in ZF set theory (Uzquiano, op. cit.).[18]

Fine (2005) has taken the interpretational modality to be imperatival, and has suggested that a dynamic logic might be an optimal means of formalizing the imperative to reinterpret quantifier domains. He (op. cit.) suggests, further, that the interpretational modality might be characterized as a program, whose operations can take the form of 'simple' and 'complex' postulates which enjoin subjects to reinterpret the domains. Uzquiano's (op. cit.) generalization of the interpretational modality, in order to target the reinterpretation of the extensions of terms such as the membership relation, can similarly be treated.

In propositional dynamic logic (PDL), there are an infinite number of diamonds, with the form $\langle \pi \rangle$.[19] π denotes a non-deterministic program, which in the present setting will correspond to Fine's postulates adumbrated in the foregoing. $\langle \pi \rangle \phi$ abbreviates 'some execution of π from the present state entrains a state bearing information ϕ'. The dual operator is $[\pi]\phi$, which abbreviates 'all executions of π from the present state entrain a state bearing information ϕ'. π^* is a program that executes a distinct program, π, a number of times ≥ 0. This is known as the iteration principle. PDL is similarly closed under finite and infinite unions. This is referred to as the 'choice' principle: If π_1 and π_2 are programs, then so is $\pi_1 \cup \pi_2$. The forth condition is codified by the 'composition' principle: If π_1 and π_2 are programs, then $\pi_1;\pi_2$ is a program (intuitively: the composed program first executes π_1 then π_2). The back condition is codified by Segerberg's induction axiom (Blackburn et al., op. cit: p. 13): All executions of π^* (at t) entrain

[18]A variant strategy is pursued by Eagle (2008). Eagle suggests that the relation between rational intuition and conceptual elucidation might be witnessed via associating the fundamental properties of the entities at issue with their Ramsey sentences; i.e., existentially generalized formulas, where the theoretical terms therein are replaced by second-order variables bound by the quantifiers. However, the proposal would have to be expanded upon, if it were to accommodate Gödel's claim that mathematical intuitions possess a modal profile.

[19]Cf. Blackburn et al., op. cit.: 12-14. A semantics and proof-theory for PDL are outlined in Hoare (1969); Pratt (1976); Goldblatt (1987: ch. 10; 1993: ch. 7) and van Benthem (2010: 158).

the following conditional state: If it is the case that (i) if ϕ, then all the executions of π (at t) yield ϕ; then (ii) if ϕ, then all executions of π^* (at t) yield ϕ. Formally, $[\pi^*](\phi \to [\pi]\phi) \to (\phi \to [\pi^*]\phi)$.

We augment, then, the provability logic for Gödelian intuition with the axiom, $\Box\phi \to [\pi]\phi$, in order to yield a bimodal, dynamic provability logic thereof.

Crucially, the iteration principle permits π^* to be interpreted as a fixed point for the equation: x $\iff \phi \lor \diamond$x. The smallest solution to the equation will be the least fixed point, μx.$\phi \lor \diamond$x, while the largest solution to the equation, when $\pi^* \lor \infty_\diamond$, will be the greatest fixed point, νx.$\phi \lor \diamond$x. Janin and Walukiewicz (op. cit.) have proven that the modal μ-calculus is equivalent to the bisimulation-invariant fragment of second-order logic.

Fine's simple dynamic-postulational modality takes, then, the form:

'(i) Introduction. !x.C(x)', which states the imperative to: '[I]ntroduce an object x [to the domain] conforming to the condition C(x)'.

Fine's complex dynamic-postulational modalities are the following:

(ii) 'Composition. Where β and γ are postulates, then so is β;γ. We may read β;γ as: do β and then do γ; and β;γ is to be executed by first executing β and then executing γ.
(iii) Conditional. Where β is a postulate and A an indicative sentence, then A $\to \beta$ is a postulate. We may read A $\to \beta$ as: if A then do β. How A $\to \beta$ is executed depends upon whether or not A is true: if A is true, A $\to \beta$ is executed by executing β; if A is false, then A $\to \beta$ is executed by doing nothing.
(iv) Universal. Where β(x) is a postulate, then so is $\forall$xβ(x). We may read $\forall$xβ(x) as: do β(x) for each x; and $\forall$xβ(x) is executed by simultaneously executing each of β(x1), β(x2), β(x3), ..., where x1, x2, x3, ... are the values of x (within the current domain). Similarly for the postulate $\forall$Fβ(F), where F is a second-order variable.
(v) Iterative Postulates. Where β is a postulate, then so is β^*. We may read β^* as: iterate β; and β^* is executed by executing β, then executing β again, and so on ad infinitum' (op. cit.: 91-92).

Whereas Fine avails of the foregoing interpretational modalities in order both to account for the notion of indefinite extensiblity and to demonstrate

how unrestricted quantification can be innocuous without foundering upon Russell's paradox (op. cit.: 26-30), the primary interest in adopting modal μ provability logic as the logic of rational intuition is its capacity to account for reinterpretations of mathematical vocabulary and quantifier domains; and thus to illuminate how the precise mechanisms codifying modalized rational intuition might be able to entrain advances in conceptual elucidation.

Finally, the computational profile of modalized rational intuition can be outlined as follows. In category theory, a category C is comprised of a class Ob(C) of objects a family of arrows for each pair of objects C(A,B) (Venema, 2007: 421). An **E**-coalgebra is a pair $\mathbb{A} = (A, \mu)$, with A an object of C referred to as the carrier of $\mathbb{A}$, and μ: A $\to$ **E**(A) is an arrow in C, referred to as the transition map of $\mathbb{A}$ (390). A coalgebraic model of deterministic automata can be thus defined (391). An automaton is a tuple, $\mathbb{A} = \langle$A, a_I, C, δ, F$\rangle$, such that A is the state space of the automaton $\mathbb{A}$; $a_I \in$A is the automaton's initial state; C is the coding for the automaton's alphabet, mapping numerals to properties of the natural numbers; δ: A X C $\to$ A is a transition function, and F$\subseteq$ A is the collection of admissible states, where F maps A to $\{1,0\}$, such that F: A $\to$ 1 if a$\in$F and A $\to$ 0 if a$\notin$F (op. cit.). The modal profile of coalgebraic automata can be witnessed both by construing the transition function as a counterfactual conditional (cf. Stalnaker, 1968; Williamson, 2007), and in virtue of the convergence of coalgebraic categories of automata with coalgebraic models of modal logic (407). Let

$\diamond\phi \equiv \nabla\{\phi, \text{T}\}$,

$\square\phi \equiv \nabla\varnothing \vee \nabla\phi$ (op. cit.).

Let an **E**-coalgebraic modal model, $\mathbb{A} = \langle$S,λ,R[.]$\rangle$, such that $\mathbb{S}$,s $\Vdash \nabla\Phi$ if and only if, for all (some) successors σ of s$\in$S, $[\Phi,\sigma(s)\in$**E**$(\Vdash_{\mathbb{A}})]$ (op. cit.). $\mathbb{A}$ is then the category of modal coalgebraic automata.

The philosophical significance of the foregoing is that the modal logic of rational intuition can be interpretable in the category of modal coalgebraic automata. The foregoing accounts for the distinctively computational nature of the modal profile of rational intuition.

13.4   Concluding Remarks

In this note, I have endeavored to outline the modal logic of Gödel's conception of intuition, in order both (i) to provide a formally tractable foundation thereof, and thus (ii) to answer the primary objection to the notion as a

viable approach to the epistemology of mathematics. I have been less concerned with providing a defense of the general approach from the array of objections that have been proffered in the literature. Rather, I have sought to demonstrate how the mechanisms of rational intuition can be formally codified and thereby placed on a secure basis.

Among, e.g., the most notable remaining objections, Koellner (2009) has argued that the best candidates for satisfying Gödel's conditions on being intrinsically justified are reflection principles, which state that the height of the hierarchy of sets, V, cannot be constructed 'from below', because, for all true formulas in V, the formulas will be true in a proper initial segment of the hierarchy. Koellner's limitative results are, then, to the effect that reflection principles cannot exceed second-order universal quantification without entraining inconsistency (op. cit.).[20] Another crucial objection is that the properties of rational intuition, as a species of cognitive phenomenology, lack clear and principled criteria of individuation. Burgess (2014) notes, e.g., that the role of rational intuition in alighting upon mathematical truths might be distinct from the functions belonging to what he terms a 'heuristic' type of intuition. The constitutive role of the latter might be to guide a mathematician's non-algorithmic insight as she pursues an informal proof. A similar objection is advanced in Cappelen (2012: 3.2-3.3), who argues that – by contrast to the properties picked out by theoretical terms such as 'utility function' – terms purporting to designate cognitive phenomenal properties both lack paradigmatic criteria of individuation and must thereby be a topic of disagreement, in virtue of the breadth of variation in the roles that the notion has been intended to satisfy.

The foregoing issues notwithstanding, I have endeavored to demonstrate that – as with the property of knowledge – an approach to the notion of intuition-that which construes the notion as a modal operator, and the provision thereof with a philosophically defensible logic, might be sufficient to counter the objection that the very idea of rational intuition is mysterious and constitutively unconstrained.

---

[20]The issue of accounting for third-order variables in order to state, e.g., Cantor's Theorem and the Generalized Continuum Hypothesis might yet be redressed, if one were to follow Shapiro (1991: 103-105), and avail of second-order paraphrases for formulas which concern sets and sequences of sets.

Chapter 14

An Epistemicist Solution to the Alethic Paradoxes

This essay targets a series of potential issues for the discussion of, and resolution to, the alethic paradoxes advanced by Scharp, in his *Replacing Truth* (2013). I aim, then, to provide a novel, epistemicist treatment of the alethic paradoxes. In response to Curry's paradox, the novel, epistemicist solution that I advance enables the retention of both classical logic and the traditional rules for the alethic predicate: truth-elimination and truth-introduction. By availing of epistemic logic, the epistemicist approach permits, further, of a precise, descriptively adequate explanation of the indeterminacy exhibited by epistemic states concerning paradoxical sentences such as the liar.

In Section **2**, Scharp's replacement strategy is outlined, and his semantic model is described in detail. Sections **3-7** examine five, crucial issues for the approach and the semantic model that Scharp proffers. The five issues target the following points of contention:

(i) Whether a positive theory of validity might be forthcoming on Scharp's approach, given that Scharp expresses sympathy with treatments on which – in virtue of Curry's paradox – validity is not identical with necessary truth-preservation;

(ii) The failure of compositionality in Scharp's Theory of Ascending and Descending Truth (ADT) and whether the theory is not, then, in tension with natural language semantics. The foregoing might be pernicious, given Scharp's use of consistency with natural language semantics as a condition for the success of approaches to the paradoxes. A related issue concerns whether it is sufficient to redress the failure of compositionality by availing

of hybrid conditions which satisfy both Ascending and Descending Truth;

(iii) Whether ADT can generalize, in order to account for other philosophical issues that concern indeterminacy;

(iv) Whether Descending Truth and Ascending Truth can countenance the manner in which truth interacts with objectivity. It is unclear, e.g., how the theorems unique to each of Descending Truth and Ascending Truth – respectively, T-Elimination and T-Introduction – can capture distinctions between the reality of the propositions mapping to 1 in mathematical inquiry, by contrast to propositions – about humor, e.g. – whose mapping to 1 might be satisfied by more deflationary conditions; and

(v) Whether the replacement strategy in general and ADT in particular can be circumvented, in virtue of approaches to the alethic paradoxes which endeavor to resolve them by targeting constraints on the contents of propositions and the values that they signify.

Despite the critical issues that are raised for the approach, I believe that the theory of ADT can be extended in order to solve several other types of paradox, beyond the paradoxes of self-reference. Scharp disavows of a unified approach to all paradoxical phenomena. However, because the theory of Ascending and Descending Truth appears to be extendible to several paradoxes, it might repay the attempt to examine the theory's possible strengths. In Section **5**, I outline, therefore, extensions of Scharp's ADT theory to the preface paradox; to the property version of Russell's paradox in the setting of unrestricted quantification; to probabilistic self-reference; and to the sorites paradox.

Finally, Section **8** examines the alethic paradoxes in the setting of epistemic logic. My epistemicist theory avoids the series of issues adducing against the ADT approach. I outline a novel, epistemicist solution to Curry's paradox, and provide a precise explanation of the epistemic indeterminacy expressed by agent's attitudes toward the liar sentence.

14.1   Scharp's Replacement Theory

Scharp avers that two main alethic principles target the use of the predicate as a device of endorsement and as a device of rejection. When the truth predicate is governed by (T-Out), then it can be deployed in the guise of a device of endorsement. When the truth predicate is governed by (T-In), then it can be deployed in the guise of a device of rejection.

Scharp's theory aims to replace truth with two distinct concepts. His explicit maneuver is to delineate the two, smallest inconsistent subsets of alethic principles; and then to pair one of the subsets with one of the replacement concepts, and the other subset with the second replacement concept.

Thus, one replacement concept will be governed by (T-In) and not by (T-Out); and the second replacement concept will be governed by (T-Out) and not by (T-In).

Scharp refers to one of his two, preferred replacement concepts as *Descending Truth* (henceforth DT). DT is governed by (T-Out).

Scharp refers to the second of his two, preferred replacement concepts as *Ascending Truth* (henceforth AT). AT is governed by (T-In).

In his 'Syntactical Treatments of Modality, with Corollaries on Reflexion Principles and Finite Axiomatizability' (1963), Montague proved that, for any predicate $H(x)$, the following conditions on the predicate are inconsistent.

Montague's (1963) Lemma 3:

(i) All instances of $H(\phi) \to \phi$ are theorems.

(ii) All instances of $H[H(\phi) \to \phi]$ are theorems.

(iii) All instances of $H(\phi)$, where ϕ is a logical axiom, are theorems.

(iv) All instances of $H(\phi \to \psi) \to [H(\phi) \to H(\psi)]$ are theorems.

(v) Q – i.e., Robinson Arithmetic – is a subtheory.

Scharp notes that Montague's conditions target only Predicate-Elimination, and are thus apt for governing DT.

Scharp argues that (v) is necessary, in order for languages that express the theory to refer to their own sentences. Condition (i) is necessary, because it captures (T-Out). Condition (ii) is necessary, because denying iterations of DT entrains a version of the revenge paradox.

Thus, either Condition (iii) or Condition (iv) must be rejected. Condition (iii) states that all tautologies are Descending True. Condition (iv) is an instance of closure. In virtue of considerations pertaining to validity (see Section **3**), Scharp is impelled to reject (iv), s.t. DT cannot satisfy closure (151).

14.1.1   Properties of DT and AT

Scharp argues that the alethic principles, DT and AT, ought to include the following.

DT ought to satisfy:

[¬-Exc: D(¬ϕ) → ¬D(ϕ)];
[∧-Exc: D(ϕ ∧ ψ) → D(ϕ) ∧ D(ψ); and
[∨-Imb: D(ϕ) ∨ D(ψ) → D(ϕ ∨ ψ).
However, DT is not governed by:
[¬-Imb: ¬T(ϕ) → T(¬ϕ); nor by
[∨-Exc: T(ϕ ∨ ψ) → T(ϕ) ∨ T(ψ).
AT ought to satisfy:
[¬-Imb: ¬A(ϕ) → A(¬ϕ);
[∧-Exc: A(ϕ ∧ ψ) → A(ϕ) ∧ A(ψ); and
[∨-Imb: A(ϕ) ∨ A(ψ) → A(ϕ ∨ ψ).
However, AT is not governed by:
[¬-Exc: T(¬ϕ) → ¬T(ϕ)]; nor by
[∧-Imb: T(ϕ) ∧ T(ψ) → T(ϕ ∧ ψ).
Scharp argues, further:
–that classical tautologies are Descending True;
–that the axioms governing the syntax of the theory are Descending True;
–that the axioms of PA are Descending True, in order to induce self-reference via Gödel-numbering; and
–that the axioms of the theories for both AT and DT are themselves Descending True (152).

DT takes classical values, and, in Scharp's theory, there are no restrictions on the language's expressive resources. This is problematic, because 'a' := '¬A(x)' and 'd' := '¬D(x)' can be countenanced in the language, and thereby yield contradictions:

Because A(x) is governed by (T-In), 'a' entails that A(a), although a states of itself that ¬A(x). Contradiction.

Because D(x) is governed by (T-Out), 'd' entails that replacing 'd' for x in 'd' is not a descending truth, i.e., ¬D(d)]. So – by condition (ii) – 'D[D(x)] → x' entails that it is not a descending truth that replacing 'd' for x in 'd' is not a descending truth [i.e., ¬D(¬d)].

Thus, Scharp concedes that there must be problematic sentences in the language for his theory, s.t. both the sentences and their negations are Ascending True, and s.t. the sentences and their negations are not Descending True (op. cit.).

Scharp endeavors to block the foregoing, by suggesting that DT can be governed by both unrestricted (T-Out), as well as a restricted version of (T-In). Similarly, AT can be governed by both unrestricted (T-In), as well as a restricted version of (T-Out).

To induce the foregoing, Scharp introduces a 'Safety' predicate, S(x). A sentence ϕ is *safe* if and only if ϕ is either (DT and AT) or not AT.

Thus,

S(ϕ) $\wedge$ ϕ $\rightarrow$ D(ϕ); and

S(ϕ) $\wedge$ A(ϕ) $\rightarrow$ ϕ.

A sentence ϕ is *unsafe* if and only if ϕ is AT and not DT:

S(ϕ) $\iff$ D(ϕ) $\vee$ ¬A(ϕ).

From which it follows that:

¬S(ϕ) $\iff$ ¬D(ϕ) $\wedge$ A(ϕ), s.t.

D(ϕ) $\rightarrow$ A(ϕ);

¬$\exists\phi$[D(ϕ)$\wedge$ ¬A(ϕ)]

A(ϕ) $\rightarrow$ ¬D(ϕ); and

¬A(ϕ) $\rightarrow$ ¬D(ϕ) (153).

Scharp avers too that AT and DT are duals. Thus,

D(ϕ) $\iff$ ¬A¬(ϕ); and

A(ϕ) $\iff$ ¬D¬(ϕ) (152).

14.1.2 Scharp's Theory: ADT/Xeno Semantics

Scharp's Theory is referred to as ADT. The necessary principles comprising ADT are as follows (cf. 154):

- Descending Truth

(D1): D(ϕ) $\rightarrow$ ϕ
(D2): D(¬ϕ) $\rightarrow$ ¬D(ϕ)
(D3): D(ϕ $\wedge$ ψ) $\rightarrow$ [D(ϕ) $\wedge$ D(ψ)]
(D4): [D(ϕ) $\vee$ D(ψ)] $\rightarrow$ D(ϕ $\vee$ ψ)
(D5): If ϕ is a classical tautology, then D(ϕ)
(D6): If ϕ is a theorem of PA, then D(ϕ)
(D7): If ϕ is an axiom of ADT, then D(ϕ).

- Ascending Truth

(A1): ϕ $\rightarrow$ A(ϕ)
(A2): ¬A(ϕ) $\rightarrow$ A(¬ϕ)
(A3): [A(ϕ) $\vee$ A(ψ)] $\rightarrow$ A(ϕ $\vee$ ψ)
(A4): A(ϕ $\wedge$ ψ) $\rightarrow$ [A(ϕ) $\wedge$ A(ψ)]
(A5): If ϕ maps to the falsum constant, then ¬A(ϕ)
(A6): If ϕ negates an axiom of PA, then ¬A(ϕ)

- Transformation Rules

(M1): D(ϕ) $\iff$ ¬A¬(ϕ)
(M2): S(ϕ) $\iff$ D(ϕ) ∨ ¬A(ϕ)
(M3): S(ϕ) ∧ ϕ → D(ϕ)
(M4): A(ϕ) ∧ S(ϕ) → ϕ
(E1): If s and t are terms; s = t; and replacing s with t in a sentence p yields a sentence q; then D(p) $\iff$ D(q)
(E2): If s and t are terms; s = t; and replacing s with t in a sentence p yields a sentence q; then A(p) $\iff$ A(q)
(E3): If s and t are terms; s = t; and replacing s with t in a sentence p yields a sentence q; then S(p) $\iff$ S(q).

14.1.3 Semantics for ADT

Scharp advances a combination of relational semantics for a non-normal modal logic, as augmented by a neighborhood semantics. (A modal logic is normal if and only if it includes axiom K and the rule of Necessitation; respectively '$\Box[\phi \to \psi] \to [\Box\phi \to \Box\psi]$' and '$\vdash\phi \to \vdash\Box\phi$'.) He refers to this as *xeno semantics.*

A model, M, of ADT is a tuple, $\langle$D, W, R, I$\rangle$, where D is a non-empty domain of entities constant across worlds, W denotes the space of worlds, R denotes a relation of accessibility on W, and I is an interpretation-function mapping subsets of D to W. The clauses for defining truth in a world in the model are familiar:

$\langle$M,w$\rangle$ $\Vdash$ ϕ iff w$\in V(\phi)$
$\langle$M,w$\rangle$ $\Vdash$ ¬ϕ iff it is not the case that $\langle$M,w$\rangle$ $\Vdash$ ϕ
$\langle$M,w$\rangle$ $\Vdash$ ϕ ∧ ψ iff $\langle$M,w$\rangle$ $\Vdash$ ϕ and $\langle$M,w$\rangle$ $\Vdash$ ψ
$\langle$M,w$\rangle$ $\Vdash$ ϕ ∨ ψ iff $\langle$M,w$\rangle$ $\Vdash$ ϕ or $\langle$M,w$\rangle$ $\Vdash$ ψ
$\langle$M,w$\rangle$ $\Vdash$ ϕ → ψ iff, if $\langle$M,w$\rangle$ $\Vdash$ ϕ, then $\langle$M,w$\rangle$ $\Vdash$ ψ
$\langle$M,w$\rangle$ $\Vdash$ ϕ $\iff$ ψ iff [$\langle$M,w$\rangle$ $\Vdash$ ϕ iff $\langle$M,w$\rangle$ $\Vdash$ ψ]
$\langle$M,w$\rangle$ $\Vdash$ $\Box\phi$ iff ∀w'[If R(w,w'), then $\langle$M,w'$\rangle$ $\Vdash$ ϕ]
$\langle$M,w$\rangle$ $\Vdash$ $\Diamond\phi$ iff ∃w'[R(w,w') and $\langle$M,w'$\rangle$ $\Vdash$ ϕ]

14.1.4 Further Montagovian Desiderata

Problem 1: Montague (op. cit.) proved that the predicate version of axiom K is inconsistent with (D1): D(ϕ) → ϕ; (D5): If ϕ is a classical tautology,

then D(ϕ); and (D6): If ϕ is a theorem of PA, then D(ϕ).

Scharp includes (D7) (159): If ϕ is an axiom of ADT, then D(ϕ). However, while Montague was one of the most precocious philosophers of the twentieth century, I suspect that Scharp does not mean to be attributing to Montague an anticipation of his theory.

Problem 2: Montague (op. cit.) proved that the predicate version of Necessitation is inconsistent with (D1): D(ϕ) $\rightarrow$ ϕ, and (D6): If ϕ is a theorem of PA, then D(ϕ).

Problem 3: The third problem is supposed to concern whether the quantifier domain is constant or variable. Scharp opts for constant-domain models without argument, so it is unclear precisely what about the quantifier domain he is targeting as problematic. (Cf. 6.7.7: 172.)

Problem 4: The fourth problem concerns the status of Scharp's predicates, AT and DT, in the relational semantics, given that the modalities in the logic for the semantics are expressed by operators.

14.1.5   Satisfying the Desiderata

To address Problems 1 and 2, Scharp augments his relational semantics with a neighborhood semantics. M = $\langle$D, W, R, I$\rangle$ is thus enriched with a neighborhood function, N, which maps sets of subsets of W to each world in W.

Necessity takes then the revised clause:

$\langle$M,w$\rangle$ ⊩ $\Box\phi$ iff $\exists X \in N(w) \forall w'[\langle$M,w'$\rangle$ ⊩ $\phi \iff w' \in X]$

Possibility takes the revised clause:

$\langle$M,w$\rangle$ ⊩ $\Diamond\phi$ iff $\neg[\exists X \in N(w) \forall w'[\langle$M,w'$\rangle$ ⊩ $\neg\phi \iff w' \in X]]$

K and Necessitation need not hold on the neighborhood structure.

A counterexample to K in the neighborhood structure (160): Let M_k be a neighborhood model, s.t. $W_k = \{a,b,c,d\}$; $N_k(a) = \{\{a,b\}, \{a,c,d\}\}$; $N_k(b) = \{b\}$; $N_k(c) = \{c\}$; $N_k(d) = \{d\}$; $V_k(\phi) = \{a,b\}$; and $V_k(\psi) = \{a,c\}$. Thus, $\langle M_k,w \rangle$ ⊩ $\Box\phi$; $\langle M_k,w \rangle$ ⊩ $\Box[\phi \rightarrow \psi]$; but it is not the case that $\langle M_k,w \rangle$ ⊩ $\Box\psi$.

So, K need not hold on the neighborhood structure. So, Montague's proof that the predicate version of K is inconsistent with Montague's conditions on predicates (interpreted, by Scharp, as expressing Descending Truth) is not pernicious.

A counterexample to Necessitation in the neighborhood structure (op. cit.): Let M_n be a neighborhood model, s.t. $W_n = \{a,b\}$; $N_n(a) = \{\{a\}\}$; $N_n(b) = \{\{b\}\}$; and $V_n(\phi) = \{a,b\}$. Thus, ϕ can be true in W, while $\Box\phi$ is false in $N_n(a)$.

So, Necessitation need not hold on the neighborhood structure. So, Montague's proof that the predicate version of Necessitation is inconsistent with the the conditions on Descending Truth is not pernicious.

So, Problems 1 and 2 can be answered.

However, the following rule is valid on neighborhood frames, and it entrains a fifth problem:

(L) If $\vdash \phi \iff \psi$, then $\vdash \Box\phi \iff \Box\psi$.

Problem 5:

The problem with regard to replacing modal operators with Descending Truth predicates hinges on (L).

Scharp writes, e.g., that Gödel's (1931) Diagonalization Lemma 'guarantees that if our language can express Peano Arithmetic or its own theory of syntax [. . .] then it will have a sentence d, s.t. ¬D(d) is provably equivalent to d' (161).

Scharp concedes that Gödel's incompleteness theorem would yield that $\vdash_{ADT}$ ¬D(d). He argues for this as follows: Suppose D(d), so if D(d) then d; but – by the Lemma – if d, then ¬D(d); so if D(d), then ¬D(d).

Similarly, take another sentence in ADT, '0 = 0', then $\vdash_{ADT}$ ¬D(d) and $\vdash$'0 = 0'. However, assuming (L) and replacing the operators therein with predicates entails that $\vdash_{ADT}$ D(d) $\iff$ D['0 = 0']. Given that it is true in ADT that D['0 = 0'], then $\vdash_{ADT}$ D(d).

Scharp argues that Problem 5 can be resolved by availing of the structure built into his xeno semantics.

Let L be a language with Boolean connectives, and the operators $\Box$, $\diamond$, and Σ. $\Box$ is the Descending Truth operator. $\diamond$ is the Ascending Truth operator. Σ is the Safety operator. A *xeno model* M = $\langle$F,R,N,V$\rangle$ where F denotes a xeno frame, R is an accessibility relation on wff in L, N is a function from W to 2^{2^w}, and V is an assignment-function from wff in L to the values [0,1].

Truth in a world is defined inductively as above.

The operators take the following clauses:

Descending Truth:

$\langle$M,w$\rangle \Vdash \Box\phi$ iff $\forall$w'$\in$W[R$_\phi$(w,w') $\to \exists X \in$N(w')$\forall$v$\in$W[$\langle$M,v$\rangle \Vdash \phi \iff$ v$\in X$]

Ascending Truth:

P(ϕ) denotes the neighborhood structure – i.e., the set of subsets of worlds – at which ϕ is true.

$\langle$M,w$\rangle \Vdash \diamond\phi$ iff ¬[$\forall$w'$\in$W[R$_{\neg\phi}$(w,w') $\to$ P(¬ϕ)$\in$N(w')]

Safety:

$\langle$M,w$\rangle \Vdash \Sigma\phi$ iff $\forall$w'$\in$W[R$_\phi$(w,w') $\rightarrow$ P(ϕ)$\in$N(w')] $\vee$ $\exists$w'$\in$W[R$_{\neg\phi}$(w,w') $\wedge$ P($\neg\phi$)$\neg\in$N(w')]

A reflexive and co-reflexive xeno frame is equivalent to a neighborhood frame:

(Reflexivity) $\forall\phi\forall$w$\in$W[R$_\phi$(w,w)] $\wedge$

(Co-reflexivity) $\forall\phi\forall$w$\in$W$\forall$w'$\in$W[R$_\phi$(w,w') $\rightarrow$ w = w']

In order to invalidate (L), Scharp constructs the following xeno model, M.

W = {a,b}

N(a) = {{a,b}}

N(b) = {{b}}

R$_S$ = {(a,a),(b,b)}

R$_T$ = {(a,a),(b,b),(a,b)}

P(S) = {(a,b)}

P(T) = {(a,b)}

$\forall$w$\in$W $\langle$M,w$\rangle \Vdash$ S iff $\langle$M,w$\rangle \Vdash$ T. So, S $\iff$ T is valid in M.

However, $\forall$w'[R$_S$(a,w')] $\rightarrow$ P(S)$\in$N(w'), so $\langle$M,w$\rangle \Vdash \neg\Box$T.

So, (L) is invalid in M, and Problem 5 can be answered.

A sentential xeno frame is *acceptable* iff

(i) $\forall$w$\in$W N(w) $\neq$ $\emptyset$

(ii) $\forall$w$\in$W $\forall$X$\in$N(w) X $\neq$ $\emptyset$

(iii) $\forall$w$\in$W $\forall$X$\in$N(w) w$\in$X

(iv) $\forall\phi\in$L$\forall$w$\in$W[R$_\phi$(w,w)]

(v) if ϕ and ψ are of the same syntactic type, then R$_\phi$ = R$_\psi$

The interest of acceptable xeno frames is that they validate operator equivalents of (D1) and (D2) in ADT. [(D1): D(ϕ) $\rightarrow$ ϕ; (D2): D($\neg\phi$) $\rightarrow$ $\neg$D(ϕ).]

Suppose that $\langle$M,w$\rangle \Vdash \Box\phi$, so $\forall$w'$\in$W[R$_\phi$(w,w') $\rightarrow$ P(ϕ)$\in$N(w')]. By (iv), R$_\phi$(w,w), so P($\neg\phi$)$\in$N(w). By (iii), $\forall$X$\in$N(w)[w$\in$X]. So, w$\in$P($\neg\phi$). Thus, $\langle$M,w$\rangle \Vdash \Box\neg\phi$. So, it is not the case that w is in P(ϕ). By (iv), P(ϕ) is not in N(w). By (iii), $\neg[\forall$w'$\in$W[R$_\phi$(w,w') $\rightarrow$ P(ϕ)$\in$N(w'). Thus, $\langle$M,w$\rangle \Vdash \neg\Box\phi$.

Problem 3 concerned whether the domain ought to be variable or constant. Scharp prescinds from targeting any issues with regard (i) to quantification and ontology, or (ii) to the interaction between DT, AT, and empirical sentences (172). He opts for a constant-domain semantics, and thus defines constant-domain xeno models (166):

A constant-domain xeno frame is a tuple, $F = \langle W,N,R_f,D\rangle$. A constant-domain xeno model adds an interpretation-function I to F, s.t. I maps pairs of F and worlds w to subsets of D, s.t. $M = \langle F,R_M,I\rangle$.

A substitution is a function from a set of variables to elements of D. A substitution v' is *x-variant* of v, if $v(y) = v'(y)$ for all variables y.

Thus,

$\langle M,w\rangle \Vdash_v F[(a_1), \ldots, F(a_m)]$, where

a_i is either an individual constant or variable iff $\langle f(a_1), \ldots, f(a_m)\rangle \in I(F,w)$, s.t. if a_i is a variable x_i, then $f(a_i) = v(x_i)$, and if a_i is an individual constant c_i, then $f(a_i) = I(c_i)$

$\langle M,w\rangle \Vdash_v \neg\phi$ iff it is not the case that $\langle M,w\rangle \Vdash_v \phi$

$\langle M,w\rangle \Vdash_v \phi \wedge \psi$ iff $\langle M,w\rangle \Vdash_v \phi$ and $\langle M,w\rangle \Vdash_v \psi$

$\langle M,w\rangle \Vdash_v \phi \vee \psi$ iff $\langle M,w\rangle \Vdash_v \phi$ or $\langle M,w\rangle \Vdash_v \psi$

$\langle M,w\rangle \Vdash_v \phi \rightarrow \psi$ iff, if $\langle M,w\rangle \Vdash_v \phi$ then $\langle M,w\rangle \Vdash_v \psi$

$\langle M,w\rangle \Vdash_v \phi \iff \psi$ iff $[\langle M,w\rangle \Vdash_v \phi$ iff $\langle M,w\rangle \Vdash_v \psi]$

$\langle M,w\rangle \Vdash_v \forall x[\phi(x)]$ iff for all x-variant v' $\langle M,w\rangle \Vdash_{v'} \phi(x)$

$\langle M,w\rangle \Vdash_v \exists x[\phi(x)]$ iff for some x-variant v' $\langle M,w\rangle \Vdash_{v'} \phi(x)$

$\langle M,w\rangle \Vdash \Box\phi$ iff $\forall w'\in W[R_\phi(w,w') \rightarrow \exists X\in N(w')\forall v\in W[\langle M,w\rangle \Vdash [\phi \iff v\in X]$

$\langle M,w\rangle \Vdash \Diamond\phi$ iff $\exists w'\in W[R_{\neg\phi}(w,w') \wedge P(\neg\phi)$ is not in $N(w')$

14.2 Issue 1: Validity

Scharp mentions Field's (2008) argument against identifying validity with necessary truth-preservation, although does not reconstruct the argument.

In order to argue against identifying validity with necessary truth-preservation, Field draws, inter alia, on Curry's Paradox.

The argument from Curry's Paradox is such that – by (T-In) and (T-Out) – one can derive the following. If ϕ is a false sentence then,

1. $\phi \iff [T(\phi) \rightarrow \bot]$
2. $T(\phi) \iff [T(\phi) \rightarrow \bot]$
3. $T(\phi) \rightarrow [T(\phi) \rightarrow \bot]$
4. $[T(\phi) \wedge T(\phi)] \rightarrow \bot$ (by importation)
5. $T(\phi) \rightarrow \bot$
6. $[T(\phi) \rightarrow \bot] \rightarrow T(\phi)$
7. $T(\phi)$
8. $\bot$

So, necessary truth-preservation entails contradiction.

However, the argument need not be valid, if one preserves (T-In) and (T-Out) yet weakens the logic. One can avail of the strong Kleene valuation scheme, such that $|\phi|$ is ungrounded, i.e. maps to 1/2. One can then add a Determinacy operator, such that it is not determinately true that ϕ and it is not determinately true that not ϕ; so, it is indeterminate whether ϕ.

Field argues, in virtue of the foregoing, that validity ought to be a primitive. In more recent work, Field (2015) argues that validity is primitive if and only if it is 'genuine', such that the notion cannot be identical with either its model-theoretic or proof-theoretic analyses. As an elucidation of the genuine concept, he writes that 'to regard an inference or argument as valid is to accept a constraint on belief [. . . ; s.t.] (in the objective sense of 'shouldn't') we shouldn't fully believe the premises without fully believing the conclusion' (op. cit.). (The primitivist notion is intended to hold, as well, for partial belief.)

Scharp is persuaded by Field's argument, and endorses, in turn, a primitivist notion of validity, as a primitive canon of reasoning without necessary truth-preservation. Scharp takes this to be sufficient for the retention of Condition (iii), in Montague's Lemma (151). Scharp does not provide any further account of the nature of validity in the book. In later sections of the book, he reiterates his sympathy with Field's analysis, and also avails of Kreisel's 'squeezing' argument (section 8.8), to the effect that the primitive notion of validity extensionally coincides with a formal notion of validity (i.e., derivation in a first-order axiomatizable quantified logic with identity). However, one potential issue is that, in a subsequent passage, Scharp writes that: 'an argument whose premises are the members of the set G and whose conclusion is p is valid iff for every point of evaluation e [i.e., index], if all members of G are assigned tM-value [i.e., an AT- or DT-value of] 1 at e, then p is assigned tM-value 1 at e' (240); and this would appear to be a definition of validity as necessary truth-preservation.

The primitivist approach to validity is the primary consideration that Scharp explicitly avails of, when arguing that closure ought to be rejected (Condition iv, in Montague's Lemma), rather than rejecting logical tautologies as candidates for the axioms of ADT (Condition iii, in Montague's Lemma) (151). So, further remarks about the nature of validity would have been welcome. An objection to prescinding from more substantial remarks about the nature of validity might also be that Scharp exploits claims with regard to its uses. So, e.g., he writes that 'a valid argument will never take

one from descending truths to something not ascending true' (177). However, that claim is itself neither a consequence of either Kreisel's squeezing argument, nor the primitivist approach to validity.

14.3 Issue 2: Hybrid Principles and Compositionality

- $\wedge$-T-Imb.

  $D(\phi) \wedge D(\psi) \rightarrow A(\phi \wedge \psi)$

  v.

  $T(\phi) \wedge T(\psi) \rightarrow T(\phi \wedge \psi)$

- $\vee$-T-Exc.

  $D(\phi \vee \psi) \rightarrow A(\phi) \vee A(\psi)$

  v.

  $T(\phi \vee \psi) \rightarrow T(\phi) \vee T(\psi)$

  (cf. 147, 171)

Feferman's (1984) theory countenances a primitive truth predicate (Feferman-true, in what follows); a primitive falsity predicate; as well as a Determinacy operator (op. cit.). This is by salient contrast to Scharp's approach, on which truth is replaced with DT and AT. Scharp argues that Feferman overemphasizes the significance of the compositionality of his Determinacy operator, at the cost of not having either logical truths or the axioms of his own theory satisfy Feferman-truth. By contrast, Scharp believes that he can avail of hybrid principles, such that it is not a requirement of ADT that Descending Truth and Ascending Truth obey compositionality (157).

One objection to this maneuver is that AT and DT are separated, in the hybrid principles, between the antecendent and consequent of the conditional.[1] So, it is unclear whether Scharp's hybrid principles are sufficient to redress the failure of compositionality in ADT; i.e., there being truth-conditions for sentences whose component semantic values are, alternatively, DT and AT.

---

[1] Thanks here to Stephen Read.

A further objection is that the foregoing might be in tension with Scharp's repeated mention of natural-language semantics, in order to argue against competing proposals. If natural-language semantics were to vindicate principles of compositionality, then this would provide a challenge to the empirical adequacy of Scharp's ADT theory, and thereby the viability of his replacement concepts for the traditional alethic predicate.

14.4   Issue 3: ADT and Indeterminacy

This issue concerns whether ADT might generalize, in order to account for other philosophical issues that concern indeterminacy. Whether ADT can be so extended to other issues, such as vagueness and types of indeterminacy, is not a necessary condition on the success of the theory. However, it might be a theoretical virtue of other accounts – e.g., classical, paracomplete, intuitionist, and supervaluational approaches – that they do so generalize; and the extensions of logic and semantics to issues in metaphysics are both familiar and legion.[2]

E.g., McGee (1991) suggests replacing the truth predicate with (i) a vague truth predicate, and (ii) super-truth. The replacement predicates are not intended for deployment in inferences implicated in reasoning, such as conditional proof and arguments by reductio (155). McGee introduces a Definiteness operator, μ, in order to yield the notion of super-truth relative to a set of precisifications. There is thus a truth predicate and a super-truth predicate. Super-truth is governed by (T-In) and (T-Out). Vague truth is governed by neither.

Thus:

–If μ(p), then μ'T(p)'

(If p is definitely true, then 'p is true' is definitely true)

–If μ(¬p), then μ¬'T(p)'

(If p is definitely not true, then 'p is true' is definitely not true)

–If p is vague, then 'T(p)' is vague

(vagueness here is secured by availing of the strong Kleene valuation scheme, such that p is ungrounded, i.e. maps neither to true nor false, and rather to .5)

McGee endeavors to avoid Revenge, by arguing that

---

[2]Cf. Williamson (forthcoming), for an argument for the retention of classical logic despite the semantic paradoxes, based on the abductive strength of its generalization.

'u' := 'u is false or u is vague'

collapses to u is vague. So, u is not definitely true, and not definitely vague. Further, u is not derivable within McGee's supervaluationist theory, nor within a separate, fixed-point theory that he also advances.

Scharp raises several issues for the supervaluational approach. One issue is that vague sentences cannot be precisified via supervaluation – i.e., rendered determinately true – on pain of Revenge (156).

Scharp argues that Descending Truth and Ascending Truth obey (T-Out) and (T-In), respectively, whereas – according to McGee – vague sentences do not. So, McGee's replacement restricts expressivity, whereas Scharp argues that there are no expressive restrictions on his proposal. Scharp notes, as well, that some of the axioms for McGee's theory are not definitely true – and are thus vague and not governed by T-Out or T-In – which would appear to be a considerable objection.

However, Field's (2008) approach – K3 plus a Determinacy operator, with a multi-valued semantics for the conditional – would appear to remain a viable proposal. Extensions of Field's proposal can be to an explanation of vagueness (Field, op. cit: ch. 5); to the logic of doxastic states (cf. Caie, 2012); and to the model-theory of metaphysics.

With regard, e.g., to the extension of Field's treatment of the paradoxes to the logic of doxastic states, Caie demonstrates that – rather than rejecting the Liar sentence – it would no longer be the case, by K3 and indeterminacy, that one could believe the Liar, and it would no longer be the case that one ought not to believe the Liar. In section **8**, I provide an epistemicist approach to Curry's paradox and the Liar which is able to retain both classical logic and the normal truth rules.

With regard to the extension of Field's treatment of the paradoxes to the logic and model-theory of metaphysics, consider the following. Given Curry's paradox, the validity of an epistemic norm might depend, for its explanation, on one's choice of logic. However, one's choice of logic might depend for its explanation on considerations from metaphysics. Suppose, e.g., that one distinguishes between fundamental and derivative metaphysical states of affairs. The fundamental states of affairs might concern the entities located in $3n$-dimensional spacetime, such as whatever is represented by the wavefunction. The derivative states of affairs might concern emergent entities located in lower, 3-dimensional spacetime, such as the clouds. In order to capture the priority of the fundamental to the derivative, fundamental states of affairs could take the classical values, [0,1]; by contrast, derivative states

of affairs could take the value .5 (in K3+indeterminacy), such that – while fundamental states are always either true or false – it is not determinate that a derivative state of affairs obtains, and it is not determinate that a derivative state of affairs does not obtain.

On the supervaluational treatment of the paradoxes, the approach can more generally be extended in order to account, e.g., for the metaphysical issues surrounding fission cases and indeterminate survival. Approaches which avail of a supervaluational response to fission scenarios, and similar issues at the intersection of nonclassical logic, metaphysical indeterminacy, and decision theory, can be found, e.g., in Williams (2014).

Thus, while it is not a necessary condition on the success of treatments to the alethic paradoxes that their proposals can generalize – in order, e.g, to aid in the resolution of other philosophical issues such as epistemic and metaphysical indeterminacy – there are viable proposals which can be so extended. The competing approaches thus satisfy a theoretical virtue that might ultimately elude ADT.

In his discussion of Priest's (2006) inclosure schema, Scharp disavows of a unified solution to the gamut of paradoxical phenomena (Scharp, 2013: 288). Despite the foregoing, I believe that there are at least four positive extensions of Scharp's theory of Ascending and Descending Truth that he does not discuss, and yet that might merit examination.

14.4.1   First Extension: The Preface Paradox

The first extension of the theory of ADT might be to the preface paradox. A set of credence functions is *Easwaran-Fitelson-coherent* if and only if (i) the credences are governed by the Kolmogorov axioms; and it is not the case both (ii) that one's credence is dominated by a distinct credence, s.t. the distinct credence is closer to the ideal, vindicated world, while (iii) one's credence is assigned the same value as the remaining credences, s.t. they are tied for closeness (cf. Easwaran and Fitelson, 2015).[3] Rather than eschew of consistency in favor of a weaker epistemic norm such as Easwaran-Fitelson coherence, the ADT theorist might argue that consistency can be preserved,

---

[3]A credence function is here assumed to be a real variable, interpreted as a subjective probability density. The real variable is a function to the [0,1] interval, and is further governed by the Kolmogorov axioms: normality, 'Cr(T) = 1'; non-negativity, 'Cr(ϕ) $\geq$ 0'; finite additivity, 'for disjoint ϕ and ψ, Cr($\phi \cup \psi$) = Cr(ϕ) + Cr(ψ)'; and conditionalization, 'Cr($\phi|\psi$) = Cr($\phi \cap \psi$) / Cr(ψ)'.

because the preface sentence, 'All of the beliefs in my belief set are true, and one of them is false' might be Ascending True rather than Descending True. Because the models in Scharp's replacement theory can preserve consistency in response to the Preface, ADT might, then, provide a compelling alternative to the Easwaran-Fitelson proposal.

14.4.2 Second Extension: Absolute Generality

A second extension of Scharp's ADT theory might be to a central issue in the philosophy of mathematics; namely unrestricted quantification. A response to the latter might further enable the development of the property versions of AT and DT: i.e., being Ascending-True-of and being Descending-True-of. For example, Fine (2005) and Linnebo (2006) advance a distinction between sets and interpretations, where the latter are properties; and suggest that inconsistency might be avoided via a suitable restriction of the property comprehension scheme.[4] A proponent of Scharp's ADT theory might be able: (i) to adopt the distinction between extensional and intensional groups (sets and properties, respectively); yet (ii) circumvent restriction of the property comprehension scheme, if they argue that R is Ascending True-of yet not Descending True-of. The foregoing maneuver would parallel Scharp's treatment of the derivation, within ADT, of the Ascending and Descending Liars and their revenge analogues (see Section **2.1** above).

14.4.3 Third Extension: Probabilistic Self-reference

A third extension of ADT might be to a self-referential paradox in the probabilistic setting. Caie (2013) outlines a puzzle, according to which:

(1) '*' := $\neg\mathrm{CrT}(*) \geq .5$

that is, (*) says of itself that it is not the case that an agent has credence in the truth of (*) greater than or equal to .5. As an instance of the T-scheme, (1) yields: '$\mathrm{T}(*) \iff \neg\mathrm{CrT}(*) \geq .5$'. However, $\mathrm{CrT}(*)$ ought to map to the interval between .5 and 1. Then, '$\mathrm{Cr}(\phi) + \mathrm{Cr}(\neg\phi) \neq 1$', violating the normality condition which states that one's credences ought to sum to 1.

---

[4]See Field (2004; 2008) for a derivation of the Russell property, R, given the 'naive comprehension scheme: $\forall u_1 \ldots u_n \exists y [\mathrm{Property}(y) \land \forall x(x \text{ instantiates } y \iff \Theta(x, u_1 \ldots u_n))]$' (2008: 294). R denotes 'does not instantiate itself', i.e. $\forall x[x \in \mathrm{R} \iff \neg(x \in x)]$, s.t. $\mathrm{R} \in \mathrm{R} \iff \neg(\mathrm{R} \in \mathrm{R})$ (2004: 78).

In ADT, the probabilist self-referential paradox might be blocked as follows. Axiom (A2) states that $\neg A(\phi) \to A(\neg\phi)$; so if it is not an Ascending Truth that ϕ, then it is an Ascending Truth that not ϕ. However, (A2) does not hold for Descending Truth. Thus, in the instance of the T-scheme which states that 'T(*) $\iff \neg CrT(*) \geq .5$', the contraposition from '$\neg CrT(*) \geq .5$' to 'Cr($\neg T(*) \geq .5$' is Ascending True, but not Descending True. So, if the contraposition from '$\neg CrT(*)$' to 'Cr($\neg T(*)$)' is not Descending True, then the transition from 'Cr(ϕ) + $\neg$Cr(ϕ) = 1' to 'Cr(ϕ) + Cr($\neg\phi$) = 1' is not Descending True. Similarly, then, to the status of the Descending Liar in ADT, the derivation of contradiction from the probabilist self-referential sentence, (1), is Ascending True, but not Descending True.

14.4.4 Fourth Extension: The Sorites Paradox

A fourth extension of ADT might, finally, be to the sorites paradox. Scharp's xeno semanics is non-normal, such that the accessibility relation is governed by the axioms T (reflexivity) and 4 (transitivity), although not by axiom K. Suppose that there is a bounded, phenomenal continuum from orange to red, beginning with a color hue, c_i, and such that – by transitivity – if c_i is orange, then c_{i+1} is orange. The terminal color hue, in the continuum, would thereby be orange and not red. The transitivity of xeno semantics explains the generation of the sorites paradox. However, xeno semantics appears to be perfectly designed in order to block the paradox, as well: The neighborhood function in Scharp's xeno semantics for ADT is such that one can construct a model according to which transitivity does not hold. Let M_k be a neighborhood model, s.t. W_k = {a,b,c}; N_k(a) = {a,b}; N_k(b) ={a,b,c}; N_k(c) = {b,c}; $V_k(\phi)$ = {a,b}. Thus, $\langle M_k,a \rangle \Vdash \Box\phi$; but not $\langle M_k,b \rangle \Vdash \Box\phi$. So, it is not the case that $\langle M_k,a \rangle \Vdash \Box\Box\phi$; so transitivity does not hold in the model. Scharp's semantics for his ADT theory would thus appear to have the resources both to generate, and to solve, the sorites paradox.[5]

---

[5]Scharp suggests that the truth predicate is contextually invariant, although assessment-sensitive (9.4). A second means by which the proposal could be extended in order to account for vagueness is via its convergence with the interest-relative approach advanced by Fara (2000; 2008).

14.5   Issue 4: Descending Truth, Ascending Truth, and Objectivity

Scharp claims that considerations of space do not permit him to elaborate on the interaction between Descending Truth, Ascending Truth, and objectivity (Section 8.3). Suppose that – depending on the target domain of inquiry – the truth-conditions of sentences might be sensitive to the reality of the objects and properties that the sentences concern. So, e.g., second-order implicit definitions for the cardinals might be true only if the terms therein refer to abstract entities. By contrast, what is said in sentences about humor might be true, if and only if the sentence satisfies deflationary conditions such as the T-schema.

Another objection to the replacement strategy, and of Scharp's candidate replacements in particular, is that it is unclear how – in principle – either Descending Truth or Ascending Truth can be deployed in order to capture the foregoing distinctions.

14.6   Issue 5: Paradox, Sense, and Signification

One final objection concerns the general methodology of the book. Scharp proceeds by endeavoring to summarize all of the extant approaches to the alethic paradoxes in the literature, and to marshall at least one issue adducing against their favor. However, there are two approaches to the paradoxes that Scharp overlooks. The first approach targets the notion of what is said by an utterance, i.e. the properties of sense and signification that a sentence might express. One such proposal is inspired by Bradwardine (c.1320/2010) and pursued by Read (2009). According to the proposal, if a sentence such as the Liar does not wholly signify that it is true, then one invalidates T-Introduction for the sentence. In a similar vein, Rumfitt (2014) argues that paradoxical sentences are a type of Scheingedanken, i.e. mock thoughts that might have a sense, although take no value; so, T-Introduction is similarly restricted.

Scharp takes it to be a virtue of his account that he can retain the disquotational principles, even though they get subsequently divided among his replacement concepts. He might then reply to the foregoing proposal by sug-

gesting that they similarly induce expressive restrictions in a manner that his approach can circumvent.

However, there are other approaches which avail of what I shall refer to as the sense and signification strategy, and which eschew of neither T-Elimination nor T-Introduction. Modulo a semantics for the conditional, K3 and indeterminacy at all orders ensures not only that hyper-determinacy – and therefore an assignment of classical values to the paradoxical sentences – is circumvented; but, furthermore, that revenge sentences cannot be derived either. Against this approach, Scharp reiterates his concern with regard to restrictions on expression. He writes, e.g., that 'Field avoids revenge only by an expressive limitation on his language' (107). However, a virtue of the approach is that, as in xeno semantics for ADT, T-Elimination and T-Introduction are preserved. Against ADT theory, K3+indeterminacy does not arbitrarily select the alethic principles that the semantic theory should satisfy. Crucially, moreover, the approach does not say more than one should like it to, as witnessed, e.g., by the derivability in ADT of both the DT and AT Liars and their revenge analogues. Rather, the language of paracompleteness and indeterminacy demonstrates that – without the loss of the foundational principles governing the alethic predicate – there are propositions which can satisfy the values in an abductively robust semantic theory, and thereby exhibit their content, although what they show cannot ultimately be said.

14.7   Epistemicism and Alethic Paradox

Finally, the second approach that Scharp does not consider is one that has recently been developed by the present author. This approach provides an *epistemicist* solution to the alethic paradoxes, and is able to retain both classical logic and a univocal, non-replacement, alethic predicate which obeys both truth-elimination and truth-introduction. In the epistemic modal system at issue, the box-operator, $\Box$, is intereted as 'the agent knows that' and the diamond-operator, $\diamond$, is intereted as 'the agent believes that'. Belief is the dual of knowledge: 'B$\phi \iff \neg$K$\neg\phi$'. The epistemic system validates axiom T, which records the factivity of knowledge: 'K$\phi \rightarrow \phi$'. Curry's paradox is, again, the following:[6]

---

[6]Read (2010) notes that there are at least three forms of Curry's paradox. One form, attributed to Albert of Saxony, is conjunctive: for any sentence A, 'ϕ' := A $\wedge$ F(ϕ). A

For any false sentence, ϕ,

1. $\phi \iff [T(\phi) \to \bot]$
2. $T(\phi) \iff [T(\phi) \to \bot]$
3. $T(\phi) \to [T(\phi) \to \bot]$
4. $[T(\phi) \wedge T(\phi)] \to \bot$ (by importation)
5. $T(\phi) \to \bot$ (by contraction)
6. $[T(\phi) \to \bot] \to T(\phi)$
7. $T(\phi)$
8. $\bot$

Rather than weaken the logic in a type-free setting (Field, 2008); eschew of contraction (Beall and Murzi, 2012); eschew of T-introduction in virtue of defective properties of signification (Read, 2009); or retain classical logic for abductive reasons and disband of one of the truth rules (Williamson, forthcoming), the current approach argues that Curry's paradox is classicaly sound; that the normal truth rules can yet be retained; and that the paradox is problematic only becuase it exhibits an instance of epistemic indeterminacy.

The epistemic indeterminacy entrained by Curry's paradox occurs because step 5 invalidates axiom K on its epistemic interpretation: '$K(\phi \to \psi) \to (K\phi \to K\psi)$'. This provides a counter-instance to epistemic closure. To see this, let ϕ denote '$T(\phi)$' and ψ denote '$\bot$'. Then:

(*) $K(T\phi \to \bot) \to [K(T\phi) \to K(\bot)]$.

Our epistemic modal system validates reflexivity, or axiom T: '$K\phi \to \phi$'.[7] Thus, in (*), $K(\bot)$ is false, because – by reflexivity – only truths can

---

second form, attributed both to Albert of Saxony and Bradwardine, is disjunctive: for any sentence A, 'ϕ' := A $\vee$ F(ϕ). A third form, owing to Löb (1955) is such that, for any sentence A: 'ϕ' := $T(\phi) \to$ A. Read (op. cit.) argues that T-introduction is the culprit in the first three forms of the paradox. A fourth variation on Curry's paradox can be found in Beall and Murzi (2012), who replace the truth predicate with a validity predicate. A fifth form is targeted by Field (2008), and is the version at issue in this essay. This form of the paradox is such that a contradiction has a unique derivation, via the truth rules, given any *false* sentence.

[7]Whether the system validates further axioms is a question which requires separate treatment. Axioms 4 [$K\phi \to KK\phi$]'; G [$BK\phi \to KB\phi$]; 4.4 [$K(\phi \wedge BK\psi) \to K(\phi \vee \psi)$]; GL [$K[K(\phi) \to \phi] \to K(\phi)$]; and Grz [$K[K(\phi \to K\phi) \to \phi] \to \phi$], have all been proposed as plausible axioms of epistemic logic. If one follows Priest (op. cit.) in arguing that the sorites paradox and the paradoxes of self-reference have the same form, and one believes that transitivity, i.e. axiom 4 on its epistemic interpretation, is the culprit in the sorites paradox, then perhaps eschewing of axiom 4 with regard to the paradoxes of self-reference might be a viable, similarly epistemicist approach. Cf. Williamson (1994; 2002), for the

be known. However, K(ϕ) can be known, because it is an instance of T-introduction. Thus, the conditional in the consequent of (*) has itself a true antecedent and false consequent, and is thus false. Finally, the conditional in the antecedent of (*), 'K(T$\phi \to \bot$)' is true, in virtue of the proof of Curry's paradox. So the instance of K expressed in (*) has a true antecedent and false consequent, providing a counter-instance to K; and so K is not a valid axiom in our system of epistemic modal logic. Thus, what is problematic about Curry's paradox is that – despite being classically sound and entailing contradiction via the normal truth rules – at least one of the steps in its derivation is epistemically indeterminate, by invalidating axiom K.

To complete the epistemicist approach to alethic paradox, we target the epistemic attitude that one ought to bear toward the liar sentence. As mentioned, Caie (2012) generalizes the strong Kleene valuation scheme, as augmented by the determinacy operator, to the doxastic setting. In order to induce consistency in an agent's belief set, the belief operator is taken to be its own dual, 'B(ϕ) $\iff$ ¬B¬ϕ'. The operator is further governed by K [B($\phi \to \psi$) $\to$ [B(ϕ) $\to$ B(ψ)]], D '[K(ϕ) $\to$ B(ϕ)], 4 [B($\phi \to$ BBϕ), and E [¬B(ϕ) $\to$ B¬B(ϕ)]. By contrast to Caie's approach, we target epistemic rather than doxastic logic. In our epistemic logic, we avoid validating the conjunction of axioms 4 and E, because their conjunction would entail that an epistemic agent's consistent theorizing is decidable, in conflict with Gödel's second incompleteness theorem (cf. Williamson, 2009). Because we should like epistemic agent's theorizing to be consistent, we avoid validating both 4 and E. The epistemic logic validates axiom T. However, Curry's paradox provides a case of epistemic indeterminacy, by invalidating axiom K. To complete the epistemicist solution to the liar paradox, I believe that the epistemic attitude toward liar sentences is one of belief, where belief is, again, the dual of knowledge: Bϕ if and only if ¬K¬ϕ. Thus, where λ is a sentence which asserts of itself that it is false, the descriptively adequate and normatively permissible attitude to bear toward the sentence is one of belief, such that one doesn't know whether it is not true. To avoid contradiction, we eschew finally of the D axiom, which records the seriality and safety conditions, according to which one ought to believe ϕ only if ϕ is known: K$\phi \to$ Bϕ. Thus, the derivation of contradiction in Curry's paradox is epistemically indeterminate, and the attitude of believing the liar sentence comprises a case of epistemic indeterminacy, as well: for all one knows it is not the case that the

foundations of the latter proposal.

liar sentence is true.

Part V: Epistemic Modality, Intention, and Decision Theory

Chapter 15

Two-dimensional Truthmaker Semantics

15.1   Introduction

Philosophical applications of two-dimensional intensional semantics have demonstrated that an account of representation which is sensitive to an array of parameters can play a crucial role in explaining the values of linguistic expressions (Kamp, 1967; Kaplan, 1979); the role of speech acts in affecting shared contexts of information (Stalnaker, 1978; Lewis, 1980,a/1998; MacFarlane, 2005); the relationship between conceivability and metaphysical possibility (Chalmers, 1996); the limits of subjectivism in ethics (Peacocke, 2003); and the viability of modal realism (Russell, 2010).

In order to circumvent issues for the modal analysis of counterfactuals (2012,a, 2012,b), and to account for the general notion of aboutness and a subject matter (2015), a hyperintensional, 'truthmaker' semantics has recently been developed by Fine (ms,a, ms,b). In this essay, I examine the status of two-dimensional indexing in truthmaker semantics, and specify the two-dimensional profile of the grounds for the truth of a formula (Section **2.2**). I proceed, then, to outline three novel interpretations of the two-dimensional, hyperintensional framework, beyond the interpretations of multiply indexed intensional semantics that are noted above. The first interpretation provides a formal setting in which to define the distinction between fundamental and derivative truths (Section **3.1**). The second interpretation concerns the interaction between the two-dimensional profile of the verifiers

for a proposition, subjective probability, and decision theory (Section **3.2**). Finally, a third interpretation of the two-dimensional hyperintensional framework concerns the types of intentional action. I demonstrate, in particular, how multiply indexed truthmaker semantics is able to resolve a puzzle concerning the role of intention in action (Section **3.3**). Section **4** provides concluding remarks.

15.2 Two-dimensional Truthmaker Semantics

15.2.1 Intensional Semantics

In his (1979), Evans endeavors to account for the phenomenon of the contingent apriori by distinguishing between two types of modality. In free logic, closed formulas may receive a positive, classical semantic value when the terms therein have empty extensions (op. cit.: 166). Informative identity statements – such as that 'Plotinus = the author of *The Enneads*' – are thus taken to be 'epistemically equivalent' to vacuously true identity statements – e.g., 'Plotinus = Plotinus' (op. cit.: 177). The apriority of the vacuously true identity statement is thus argued to be a property of the informative identity statement, as well.[1] However, the informative identity statement is contingent. For example, it is metaphysically possible that the author of *The Enneads* is Plato, rather than Plotinus.

Evans argues that the foregoing 'superficial' type of contingency at issue is innocuous, by distinguishing it from what he refers to as a 'deep' type of contingency according to which a sentence is possibly true only if it is made true by a state of affairs (185). The distinction between the types of modality consists in that superficial contingency records the possible values of a formula when it embeds within the scope of a modal operator – e.g., possibly, x is red and possibly x is blue – whereas deep contingency records whether the formula is made true by a metaphysical state of affairs. In light of the approach to apriority which proceeds via the free-logical, epistemic equivalence of vacuous and informative identity statements, a formula may thus be apriori and yet superficially contingent.[2] Evans (op. cit.: 183-184; 2004:

---

[1]A premise in the argument is that definite descriptions are non-referring, although – in free logic – still enable the sentences in which they figure to bear a positive, classical value. See Evans (op. cit.: 167-169).

[2]Evans' approach is defined within a single space of metaphysically possible worlds.

11-12) goes further and – independently developing work in two-dimensional intensional semantics by Kamp (1967), Vlach (1973), and Segerberg (1973) – treats the actuality operator as a rigidifier, such that the value of actually ϕ determines the counterfactual value of possibly ϕ.

Two-dimensional intensional semantics provides a framework for regimenting the thought that the value of a formula relative to one parameter determines the value of the formula relative to another parameter. The semantics assigns truth-conditions to formulas, and semantic values to the formula's component terms. The conditions of the formulas and the values of their component terms are assigned relative to the array of intensional parameters. So, e.g., a term may be defined relative to a context; and the value of the term relative to the context will determine the value of the term relative to an index. For example, according to Kaplan (1979), an utterance's character is a mapping from the utterance's context of evaluation to the utterance's content. According to Stalnaker (op. cit.; 2004), having distinct functions associated with the value of an utterance provides one means of reconciling the necessity of a formula presupposed by speakers with the contingency of the values of assertions made about that formula. According to Chalmers (op. cit.), the value of a formula relative to a context, which ranges over epistemically possible worlds, determines the value of a formula relative to an index, which ranges over metaphysically possible worlds. According to Lewis (op. cit.), the context may be treated as a concrete situation ranging over individuals, times, locations, and worlds; and the index may be treated as ranging over information states relative to the context. According to MacFarlane (op. cit.), formulas may receive their value relative to a context ranging over two distinct agents; the context determines the value of an index ranging over their states of information; and the value of the formula may yet be defined relative to a third parameter ranging over the

---

However, one may define the value of a formula relative to two spaces: A space of epistemic possibilities and a space of metaphysical possibilities. By contrast to securing apriority by (i) eliding the values of informative and vacuous identity statements in a free logic within a single space of metaphysical possibilties, and then (ii) arguing that apriori identity statements are superficially contingent because possibly false, an alternative approach argues that an identity statement is contingent apriori if and only if it is (i) apriori, because the statement is necessarily true in epistemic modal space, while the statement is (ii) contingent, because possibly the statement is false in metaphysical modal space (cf. Chalmers, 2006). However, a point of convergence between the above approaches is that neither avails of the notion of deep contingency – i.e., truthmaking – in order to reconcile the apriority of a formula with the possibility of its falsehood.

states of an independent, third assessor. Finally, in decision theory, the value of a formula relative to a context, which ranges over a time, location, and agent, constrains the value of the formula relative to a first index on which a space of the agent's possible acts is built, and the latter will subsequently constrain the value of the formula relative to a second index on which a space of possible outcomes may be built.

Generally, a proposition, P, is such that P is true in a Model, M, relative to a context, c, and an index, i:

- M $\Vdash$ P iff $\forall c \exists c' [\![P]\!]^{c,c'} = 1$ and $\exists i [\![P]\!]^{c,i} = 1$.

15.2.2 Truthmaker Semantics

A hyperintensional, 'truthmaker' semantics has recently been developed by Fine (ms,a, ms,b).[3] Truthmaker semantics has been applied, in order to explain the verification-conditions which ground the truth of parts of propositions, rather than of the propositions in their entirety.[4]

Truthmaker semantics is defined over a state space, F = $\langle$S, $\sqsubset\rangle$, where S is a set of states comprising a world, and $\sqsubset$ is a parthood relation on S comprising a partial order, such that it is reflexive (a $\sqsubset$ a), anti-symmetric [(a $\sqsubset$ b) $\wedge$ $\neg$(b $\sqsubset$ a)], and transitive (a $\sqsubset$ b, b $\sqsubset$ c; a $\sqsubset$ c) (ms,a: 19).

A proposition P $\subseteq$ S is verifiable if P is non-empty, and is otherwise unverifiable (20).

s,t $\sqsubset$ S are compatible if their fusion s $\sqcup$ t is a possible state of S (op. cit.).

s overlaps t if $\exists u(u \sqsubset S)$, s.t. u $\sqsubseteq$ s and u $\sqsubseteq$ t, denoted s $\sqcap$ t (op. cit.).

A model, M, over F is a tuple, M = $\langle$F,D,V$\rangle$, where D is a domain of closed formulas (i.e. propositions), and V is an assignment function mapping propositions P$\in$D to pairs of subsets of S, {1,0}, i.e. the verifier and falsifier of P, such that $[\![P]\!]^+ = 1$ and $[\![P]\!]^- = 0$ (35).

The verification-rules in truthmaker semantics are then the following:

s $\vdash$ P if s$\in[\![P]\!]^+$

---

[3]The logic for the semantics is classical. Fine (2014) develops a truthmaker semantics for intuitionistic logic.

[4]Anticipations of the theory of partial content are advanced by van Fraassen (1969) and Angell (1989; cf. Fine, 2015). Intensional interpretations of subject matters, which anticipate Fine's hyperintensional approach, are developed in Lewis (1988/1998) and Yablo (2014).

(s verifies P, if s is a truthmaker for P i.e. if s is in P's extension);

s ⊣ P if s∈⟦P⟧⁻

(s falsifies P, if s is a falsifier for P i.e. if s is in P's anti-extension);

s ⊢ ¬P if s ⊣ P

(s verifies not P, if s falsifies P);

s ⊣ ¬P if s ⊢ P

(s falsifies not P, if s verifies P);

s ⊢ P ∧ Q if ∃t,u, t ⊢ P, u ⊢ Q, and s = t ⊓ u

(s verifies P and Q, if s is the fusion of states, t and u, t verifies P, and u verifies Q);

s ⊣ P ∧ Q if s ⊣ P or s ⊣ Q

(s falsifies P and Q, if s falsifies P or s falsifies Q);

s ⊢ P ∨ Q if s ⊢ P or s ⊢ Q

(s verifies P or Q, if s verifies P or s verifies Q);

s ⊣ P ∨ Q if ∃t,u, t ⊣ P, u ⊣ Q, and s = t ⊓ u

(s falsifies P or Q, if s is the state overlapping the states t and u, t falsifies P, and u falsifies Q);

s exactly verifies P if and only if s ⊢ P if s∈⟦P⟧;

s inexactly verifies P if and only if s ▷ P if ∃s'⊑S, s' ⊢ P; and

s loosely verifies P if and only if, ∀t, s.t. s ⊔ t, s ⊔ t ⊢ P (35-36).

Differentiated contents may be defined as follows.[5] A state s ⊑ S is differentiated only if s is the fusion of distinct parts, s.t. s = s_1 ⊔ s_2. s is thereby comprised of three parts: An initial state, s_1; an additional state, s_2; and a total state, s. The three states correspond accordingly to three contents: The initial content s_1 ⊢ P_1; the additional content, s_2 ⊢ P_2; and the total content, s ⊢ P_{1,2} (ms,b: 15).

The substraction of Q from P, P – Q, is defined such that p – q = p – (q ⊓ p) (17).

Finally, subject matters may be defined as follows.

A positive subject matter, **p**⁺, expresses a verifiable proposition, ⟦P⟧⁺ (20-21).

A negative subject matter, **p**⁻, expresses a falsifiable proposition, ⟦P⟧⁻ (21).

A comprehensive subject matter expresses the fusion of the subject matters both verified and falsified by the fusion of a number of states:

---

[5]Fine (op. cit.: 8, 12) avails of product spaces in his discussion of content and subject matter, though we continue here to work with a single space for ease of exposition.

$\mathbf{p}_{1,+,-} = \mathbf{p}_{1,+} \sqcap \mathbf{p}_{1,-} = \langle s \vdash P \wedge s \dashv P \rangle$;

$\mathbf{p}_{2,+,-} = \mathbf{p}_{2,+} \sqcap \mathbf{p}_{2,-} = \langle s \vdash P_2 \wedge s \dashv P_2 \rangle$; such that,

$\mathbf{p}_{1,2,+,-} = \mathbf{p}_{1,2,+} \sqcap \mathbf{p}_{1,2,-} = \langle s \vdash P_{1,2} \wedge s \dashv P_{1,2} \rangle$ (op. cit.).

A differentiated subject matter expresses the fusion of the subject matters that are either verified or falsified by the fusion of a number of states:

$\mathbf{p}_{1,+/-} = \mathbf{p}_{1,+} \sqcup \mathbf{p}_{1,-} = \langle s \vdash P \vee s \dashv P \rangle$;

$\mathbf{p}_{2,+,-} = \mathbf{p}_{2,+} \sqcup \mathbf{p}_{2,-} = \langle s \vdash P_2 \vee s \dashv P_2 \rangle$; such that,

$\mathbf{p}_{1,2,+/-} = \mathbf{p}_{1,2,+} \sqcup \mathbf{p}_{1,2,-} = \langle s \vdash P_{1,2} \vee s \dashv P_{1,2} \rangle$ (op. cit.).

Informally, propositions P and Q express the same subject matters, $\mathbf{p}$ and $\mathbf{q}$, when the following conditions hold:

P is exactly about Q if $\mathbf{p} = \mathbf{q}$;

P is partly about Q if $\mathbf{p}$ and $\mathbf{q}$ overlap, such that $\exists u \sqsubset S(u \vdash R)$; $\forall s_1, s_2 \sqsubseteq S$, $s_1 \vdash P$, $s_2 \vdash Q$; and $u = s_1 \sqcap s_2$, such that $R = P \cap Q$;

P is entirely about Q if $\mathbf{p} \subseteq \mathbf{q}$; and

P is about Q in its entirety if $\mathbf{p} \supseteq \mathbf{q}$ (5).

15.2.3   Two-dimensional Truthmaker Semantics

In order to account for two-dimensional indexing, we augment the model, M, with a second state space, S*, on which we define both a new parthood relation, $\sqsubset^*$, and partial function, V*, which serves to map propositions in D to pairs of subsets of S*, {1,0}, i.e. the verifier and falsifier of P, such that $[\![P]\!]^+ = 1$ and $[\![P]\!]^- = 0$. Thus, $M = \langle S, S^*, D, \sqsubset, \sqsubset^*, V, V^* \rangle$. The two-dimensional hyperintensional profile of propositions may then be recorded by defining the value of P relative to two parameters, c,i: c ranges over subsets of S, and i ranges over subsets of S*.

(*) M,s∈S,s*∈S* ⊢ P iff:
(i) $\exists c_s [\![P]\!]^{c,c} = 1$ if s∈$[\![P]\!]^+$; and
(ii) $\exists i_{s*} [\![P]\!]^{c,i} = 1$ if s*∈$[\![P]\!]^+$

(Distinct states, s,s*, from distinct state spaces, S,S*, provide a two-dimensional verification for a proposition, P, if the value of P is provided a truthmaker by s. The value of P as verified by s determines the value of P as verified by s*).

The foregoing provides a two-dimensional hyperintensional semantic framework within which to interpret the values of a proposition. In order to account

for partial contents, we define the values of subpropositional entities relative again to tuples of states from the distinct state spaces in our model:

s is a two-dimensional exact truthmaker of P if and only if (*);

s is a two-dimensional inexact truthmaker of P if and only if $\exists$s'$\sqsubset$S, s$\rightarrow$ s', s' $\vdash$ P and such that

$\exists c_{s'} [\![P]\!]^{c,c} = 1$ if s'$\in[\![P]\!]^{+}$, and
$\exists i_{s*} [\![P]\!]^{c,i} = 1$ if s*$\in[\![P]\!]^{+}$;

s is a two-dimensional loose truthmaker of P if and only if, $\exists$t, s.t. s $\sqcup$ t, s $\sqcup$ t $\vdash$ P:

$\exists c_{s\sqcup t} [\![P]\!]^{c,c} = 1$ if s'$\in[\![P]\!]^{+}$, and
$\exists i_{s*} [\![P]\!]^{c,i} = 1$ if s*$\in[\![P]\!]^{+}$.

- $[\![P]\!]^{c,i}$ is exactly about $[\![Q]\!]^{c,i}$ if $f_{1-1}[\mathbf{p}^{c,i} \iff \mathbf{q}^{c,i}]$

  (Suppose that the values of P and of Q are two-dimensionally determined, as above. Then P is exactly about Q if there is a bijection between the two-dimensionally individuated subject matters that they express);

- $[\![P]\!]^{c,i}$ is partly about $[\![Q]\!]^{c,i}$ if $\mathbf{p}$ and $\mathbf{q}$ overlap, s.t. $\exists$u$\sqsubset$S, s.t. u $\vdash$ R, and $\forall$s$_1$,s$_2\sqsubseteq$S, s$_1$ $\vdash$ P, s$_2$ $\vdash$ Q, and u = s$_1$ $\sqcap$ s$_2$ such that R^{c,c} = P $\cap$ Q. A neighborhood function, A, maps u to a state s* in i where s* $\vdash$ R^{c,i}.

- $[\![P]\!]^{c,i}$ is entirely about $[\![Q]\!]^{c,i}$ if $\mathbf{p}^{c,i} \Leftarrow \mathbf{q}^{c,i}$

  (Suppose that the values of P and of Q are two-dimensionally determined. Then P is entirely about Q if there is a surjection from the subject matter of Q onto the subject matter of P);

- $[\![P]\!]^{c,i}$ is about $[\![Q]\!]^{c,i}$ in its entirety if $\mathbf{p}^{c,i} \Rightarrow \mathbf{q}^{c,i}$

  (Suppose that the values of P and of Q are two-dimensionally determined. Then P is about Q in its entirety if there is an injection from the subject matter of P onto the subject matter of Q).

15.3   New Interpretations

The two-dimensional account of truthmaker semantics provides a general framework in which a number of interpretations of the state spaces at issue can be defined. The framework may accommodate, e.g., the so-called

'semantic' and 'metasemantic' interpretations of the framework. The semantic interpretation targets, as noted, the bearing of contextual parameters on the values of terms, and provides an account of validity relative to the spaces defined in the term's intensional parameters (cf. Kaplan, op. cit.). The metasemantic interpretation accommodates, by contrast, the update effects of contingently true assertions, with the necessary propositions of which speakers' shared information states might be comprised (cf. Stalnaker, op. cit.). The framework may further be provided an 'epistemic' interpretation, in order to countenance hyperintensional distinctions in the relations between conceivability, i.e. the space of an agent's epistemic states, and metaphysical possibility, i.e. the state space of facts (cf. Chalmers, op. cit.).[6] In this section, I advance three novel interpretations of two-dimensional semantics, as witnessed by the new relations induced by the interaction between two-dimensional indexing and hyperintensional value assignments. The three interpretations concern (i) the distinction between fundamental and derivative truths; (ii) probabilistic grounding in the setting of decision theory; and (iii) the structural contents of the types of intentional action.

15.3.1   Fundamental and Derivative Truths

The first novel interpretation concerns the distinction between fundamental and derivative truths. In the foregoing model, the value of the subject matter expressed by a proposition may be verified by states in a first space, which determine, then, whether the proposition is verified by states in a second space. Allowing the first space to be interpreted so as to range over fundamental facts and the second space to be interpreted so as to range over derivative facts permits a precise characterization of the determination relations between the fundamental and derivative grounds for a truth.

Suppose, e.g., that the fundamental facts concern the computational characterization of a subject's mental states, and let the fundamental facts comprise the first state space. Let the derivative facts concern states which verify whether the subject is consciously aware of their mental representations, and let the derivative facts comprise the second state space. Finally, let ϕ be a formula in an experimental task which expresses that there is a particular valence for the contrast-level of a stimulus.[7] The formula's having a truth-

---

[6]Cf. Author (xx) for further discussion.

[7]In experimental psychology, stimuli which represent contrast gradients are referred to

maker in the first space – where the states of which range, as noted, over the subject's psychofunctional facts – will determine whether the formula has a truthmaker in the second space – where the states of which range over the mental representations of which the subject is consciously aware. If the deployment of some attentional functions provides a necessary condition on the instantiation of phenomenal awareness, then the role of the state of the attentional function in the first space in verifying ϕ will determine whether ϕ is subsequently verified relative to the second space. Intuitively: Attending to a stimulus with a particular value will constrain whether a truthmaker can be provided for being consciously aware of the stimulus. If the computational facts at issue are fundamental, and the phenomenal facts at issue are derivative, then a precise characterization may be provided of the two-dimensional relations between the verifiers which target fundamental and derivative truths.

Note that hyperintensional truthmaker semantics is consistent with a necessitist modal ontology, according to which necessarily everything is necessarily something. An apparent tension might be thought to arise by augmenting a first-order formal language with both a truthmaker principle – to the effect that a proposition is true only if there is something which entails it $[A \rightarrow \exists x \Box[\exists y(x = y \rightarrow A)]]$ – and the principle of the necessary necessity of being – which codifies the thought that necessarily everything is necessarily something $[\Box \forall x \Box y(x = y)]$. The apparent tension is that conjoining the truthmaker principle to the necessity of being entails the necessity of truth (cf. Williamson, 2013: 400). The contingency of truth is belied by there necessarily being a (possible) object which serves as either a truthmaker or a falsifier for every proposition. However, the foregoing language is consistent with there being contingent truths; e.g., formulas comprised of non-logical vocabulary, such as that there are tigers. Williamson replies by arguing against the truthmaker principle (op. cit.). He observes, however, that a hyperintensional version of the truthmaker principle is consistent with the necessary necessity of being – $A \rightarrow \exists P(A \text{ because } P)$ – which is an instance of the principle of the priority of being to truth: namely, that for all formulas πA, $[T\pi(A) \text{ because } A]$ (400-401). Consistently, then, with the necessary necessity of being, hyperintensional truthmaker semantics is one means of interpreting the explanatory relation – A 'because' B – in the principle of the priority of being to truth.

---

as Gabor patches.

233

15.3.2 Decision Theory

A second novel interpretation of two-dimensional truthmaker semantics concerns the types of intentional action, and the interaction of the latter with decision theory. As noted in the foregoing, two-dimensional intensional semantics may be availed of in order to explain how the value of a formula relative to a context ranging over an agent and time will determine the value of the formula relative to an index ranging over a space of admissible actions made on the basis of the formula, where the value of the formula relative to the context and first index will determine the value of the formula relative to a second index, ranging over a space of outcomes.

One notable feature of the decision-theoretic interpretation is that it provides a natural setting in which to provide a gradational account of truthmaking. A proposition and its component expressions are true, just if they are verified by states in a state space, such that the state and its parts fall within the proposition's extension. In decision theory, a subject's expectation that the proposition will occur is recorded by a partial belief function, mapping the proposition to real numbers in the $\{0,1\}$ interval. The subject's desire that the proposition occurs is recorded by a utilitiy function, the quantitative values of which – e.g., 1 or 0 – express the qualitative value of the proposition's occurrence. The evidential expected utility of a proposition's occurrence is calculated as the probability of its obtaining conditional on an agent's action, as multiplied by the utility to the agent of the proposition's occurrence. The causal expected utility of the proposition's occurrence is calculated as the probability of its obtaining, conditional on both the agent's acts and background knowledge of the causal efficacy of their actions, multiplied by the utility of the proposition's occurrence.

There are three points at which a probabilistic construal of the foregoing may be defined. One point concerns the objective probability that the proposition will be verified, i.e. the chance thereof. The second point concerns subjective probability with which a subject partially believes that the proposition will obtain. A third point concerns the probability that an outcome will occur, where the space of admissible outcomes will be constrined by a subject's acts. An agent's actions will, in the third case, constrain the admissible verifiers in the space of outcomes, and thus the probability that the verifier for the proposition will obtain as an outcome.[8]

---

[8] A proponent of metaphysical indeterminacy might further suggest that the verifiers are themselves gradational; thus, rather than target the probability of a verifier's realization,

In order formally to countenance the foregoing, we define a probability measure on a state space, such that the probability measure satisfies the Kolmogorov axioms: normality [$\Pr(T) = 1$]; non-negativity [$\Pr(\phi) \geq 0$]; additivity [For disjoint ϕ and $\psi[\Pr(\phi\cup\psi) = \Pr(\phi) + \Pr(\psi)]$]; and conditionalization [$\Pr(\phi|\psi) = \Pr(\phi\cap\psi) / \Pr(\psi)$]. In order to account for the interaction between objective probability and the verification-conditions in truthmaker semantics, we avail, then, of a regularity condition in our earlier model, M, in which the assignment function, V, maps propositions P$\in$D to pairs of subsets of S, $\{1,0\}$, i.e. the verifier and falsifier of P, such that $\llbracket P \rrbracket^{+} = \{0,1\}$ and $\llbracket P \rrbracket^{-} = 1 - P$. In our gradational truthmaker semantics, a state, s, verifies a proposition, P, if the probability that s is in P's extension is greater than or equal to .5:

s $\vdash$ P if $\Pr(s\in\llbracket P \rrbracket^{+}) \geq .5$.

A state, s, falsifies a proposition P if the probability that s is in P's extension is less than .5 iff the probability that s is in P's anti-extension is greater than or equal to .5

s $\dashv$ P if $\Pr(s\in\llbracket P \rrbracket^{-}) \geq .5$

iff $\Pr(s\in\llbracket P \rrbracket^{+}) < .5$.

The subjective probability with regard to the proposition's occurrance is expressed by a probability measure satisfying the Kolmogorov axioms as defined on a second state space, i.e., a space whose points are interpreted as concerning the subject's states of information. The formal clauses for partial belief in truthmaker semantics are the same as in the foregoing, save that the probability measures express the mental states of an agent, by being defined on the space of their states of information.

Finally, the interaction between objective and subjective probability measures in hyperintensional semantics may be captured in two ways.

One concerns the interaction between the chance of a proposition's oc-

---

the proponent of metaphysical indeterminacy will suggest that a proposition P is made true only to a certain degree, such that both of the proposition's extension and anti-extension will have non-negative, real values. One objection to the foregoing account of metaphysical indeterminacy for truthmakers is, however, that the metalogic for many-valued logic is classical (cf. Williamson, 2014). A distinct approach to metaphysical indeterminacy is proffered by Barnes and Williams (2011), who argue that metaphysical indeterminacy consists in persistently unpointed models, i.e. a case in which it is unclear which among a set of worlds is actual, even upon filtering the set with precisifications. A proponent of metaphysical indeterminacy for probabilistic truthmaker semantics might then argue both that the realization of a verifier has a gradational value and that it is indeterminate which of the states which can verify a given formula is actual.

currance, the subject's partial belief that the proposition will occur, and the spaces for the subjects actions and outcomes. The formal clause for the foregoing will then be as follows:

M,s $\vdash$ $[\![P]\!]^{c(c',a,o)} > .5$,

where c ranges over the space of physical states, and a probability measure recording objective chance is defined thereon; c' ranges over the space of an agent's states of information, and the value of P relative to c' determines the value of P relative to the space of the agent's acts, a, where the latter determines the space of admissible outcomes concerning P's occurrence, o. Thus, the parameters, c',a,o possess a hyperintensional two-dimensional profile, whereas the space of physical states, c, constrains the verification of the proposition's occurrence, without determining the values of the subject's partial beliefs and their subsequently conceivable actions and outcomes.

Directly accounting for the relation between c and c' – i.e., specifying a norm on the relation between chances and credences – is the second means by which to account for how objective gradational truthmakers interact with a subject's partial beliefs about whether propositions are verified. Following Lewis (1980,b/1987), a candidate chance-credence norm may be what he refers to as the 'principal principle'.[9] The principal principle states that an agent's partial belief that a proposition will be verified, conditional on the objective chance of the proposition's occurrence and the admissible evidence, will be equal to the objective chance of the proposition's occurrence itself:

$Pr_s(P \mid ch(P) \wedge E) = ch(P)$.

15.3.3 Intentional Action

A third novel interpretation of two-dimensional intensional semantics provides a natural setting in which to delineate the structural content of the

---

[9]See Pettigrew (2012), for a justification of a generalized version of the principal principle based on Joyce's (1998) argument for probabilism. Probabilism provides an accuracy-based account of partial beliefs, defining norms on the accuracy of partial beliefs with reference only to worlds, metric ordering relations, and probability measures thereon. The proposal contrasts to pragmatic approaches, according to which a subject's probability and utility measures are derivable from a representation theorem, only if the agent's preferences with regard to a proposition's occurrence are consistent (cf. Ramsey, 1926). Probabilism states, in particular, that, if there is an ideal subjective probability measure, the ideality of which consists e.g. in its matching objective chance, then one's probability measure ought to satisfy the Kolmogorov axioms, on pain of there always being a distinct probability measure which will be metrically closer to the ideal state than one's own.

types of intentional action. For example, the mental state of intending to pursue a course of action may be categorized as falling into three types, where intending-that is treated as a modal operator defined on an agent's space of states of information. One type targets a unique structural content for the state of acting intentionally, such that an agent intends to bring it about that ϕ just if the intention satisfies a clause which mirrors that outlined in the last paragraph:

- $[\![\text{Intenton-in-Action}(\phi)]\!]_w = 1$ only if $\exists\text{w'}[\![\phi]\!]^{w',c(=t,l),a,o} = 1$.

A second type of intentional action may be recorded by a future-directed operator, such that an agent intends to ϕ only if they intend to pursue a course of action in the future, only if there is a world and a future time relative to which the agent's intention is satisfied:

- $[\![\text{Intention-for-the-future}(\phi)]\!]_w = 1$ only if $\exists\text{w'}\forall\text{t}\exists\text{t'}[\text{t}< \text{t'} \wedge [\![\phi]\!]^{w',t'} = 1]$.

Finally, a third type of intentional action concerns reference to the intention as an explanation for one's course of action. In Chapter **16**, I regiment the structural content of this type of intention as a modal operator which receives its value only if a hyperintensional grounding operator which takes scope over a proposition and an action, receives a positive semantic value. The formal properties of the grounding operator follows the presentation in Fine (2012,c), according to which the operator corresponds to relations of parthood, which satisfy the logical properties of reflexivity and permit bijections between the states at issue, namely, actions and intentional descriptions.

The varieties of subject matter, as defined in two-dimensional truthmaker semantics, can be availed of in order to enrich the present approach. Having multiple state spaces from which to define the verifiers of a proposition enables a novel solution to issues concerning the interaction between action and explanation. The third type of intentional action may be regimented, as noted, by the agent's reference to an intention as an explanation for her course of action.

The foregoing may also be availed of, in order to provide a novel solution to an issue concerning the interaction betwen involuntary and intentional action. The issue is as follows. Wittgenstein (1953/2009; 621) raises the

inquiry: 'When I raise my arm, my arm goes up. Now the problem arises: what is left over if I subtract the fact that my arm goes up from the fact that I raise my arm?' Because the arm's being raised has at least two component states, namely, the arm's going up and whatever the value of the variable state might be, the answer to Wittgenstein's inquiry is presumably that the agent's intentional action is the value of the variable state, such that a combination of one's intentional action and one's arm going up is sufficient for one's raising one's arm. The aforementioned issue with the foregoing concerns how precisely to capture the notion of partial content, which bears on the relevance of the semantics of the component states and the explanation of the unique state entrained by their combination.

Given our two-dimensional truthmaker semantics, a reply to Wittgenstein's inquiry which satisfies the above desiderata may be provided. Let W express a differentiated subject matter, whose total content is that an agent's arm is raised. W expresses the total content that an agent's arm is raised, because W is comprised of an initial content, U (that one's arm goes up), and an additional content, R (that one intends to raise one's arm).

The verifier for W may be interpreted as a two-dimensional loose truthmaker. Let c range over an agent's motor states, S. Let i range over an agent's states of information, S*. We define a modal operator for intentional action on the state space of the agent's motor actions. The value of the modal operator for intentional action is positive just if a selection function, g, is a mapping from the powerset of motor actions in S to a unique state s' in S. This specifies the initial, partial content, U, that one's arm goes up. An intention may then be defined as a unique state, s*, in the agent's state of information, S*. The state, s*, specifies the additional, partial content R, that one intends to raise one's arm.

Formally:

s ⊢ U only if ∃s'⊏S, such that g: s→ s', s.t. s' ⊢ U,

∃s*, s* ⊢ R, and

W = U ⊔ R.

The two-dimensional loose truthmaker for one's arm being raised may then be defined as follows:

$\exists c_{s \to s'} [\![W]\!]^{c,c} = 1$ if s'∈$[\![W]\!]^{+}$, and

$\exists i_{s*} [\![W]\!]^{c,i} = 1$ if s*∈$[\![W]\!]^{+}$.

Intuitively, the value of the total content that one's arm is raised is defined relative to a set of motor states – where a first intentional action selects a series of motor states which partly verify that one's arm goes up. The value

238

of one's arm being raised, relative to (the intentionally modulated) motor state of one's arm possibly going up, determines the value of one's arm being raised relative to the agent's distinct intention to raise their arm. The agent's first intention selects among the admissible motor states, and – all else being equal – the motor states will verify the fact that one's arm goes up. Recall that the value of a formula relative to a context determines the value of the formula relative to an index. As follows, the priority of the motor act to the subsequent intention to raise one's arm is thus that it must first be possible for one's arm to go up in order to determine whether the subsequent intention to raise one's arm can be satisfied.[10] The fusion of (i) the state corresponding to the initial partial content that one's arm goes up, and (ii) the state corresponding to the additional partial content that one intends to raise one's arm, is sufficient for the verification of (iii) the state corresponding to the total content that one's arm is raised.

15.4 Concluding Remarks

In this essay, I have endeavored to establish foundations for the interaction between two-dimensional indexing and hyperintensional semantics. I examined, then, the philosophical significance of the framework by developing three, novel interpretations of two-dimensional truthmaker semantics, in light of the new relations induced by the model.

The first interpretation enables a rigorous characterization of the distinction between fundamental and derivative truths. The second interpretation evinces how the elements of decision theory are definable within the two-dimensional hyperintensional setting, and a novel account was then outlined concerning the interaction between probability measures and hyperintensional grounds. The third interpretation of two-dimensional hyperintensional semantics concerns the structural content of the types of intentional action. Finally, I demonstrated how the hyperintensional array of state spaces, relative to which propositions may be verified, may serve to resolve a previously intransigent issue concerning the role of intention in action.

---

[10]The role of the first intention in acting as a selection function on the space of motor actions corresponds to the comparator functions stipulated in the contemporary cognitive science of action theory. For further discussion of the comparator model, see Frith et al. (2000) and Pacherie (2012).

Chapter 16

Epistemic Modality, Intention, and Decision Theory

Formal treatments of imperatival notions have been pursued both logically and semantically. In the logical setting, deontic claims have been interpreted as types of a modal operator, where a condition holding across the points of a space abbreviates the property of obligation, and its dual abbreviates the property of permissibility.[1] In the twentieth century, research in deontic logic has examined the validity of the rule of necessitation ($\vdash \phi \to \vdash \Box \phi$) (von Wright, 1981); modal axiom 4 ($\Box \phi \to \Box \Box \phi$) (cf. Barcan, 1966); and modal axiom GL [$\Box(\Box \phi \to \phi) \to \Box \phi$] (cf. Smiley, 1963). The semantic approach has been inspired by the works of Kratzer (1977, 2012), Stalnaker (1978), and Veltman (1996), arguing that there are modal operaters on a set of points which are not straightforwardly truth-conditional, instead recording an update on that set which is taken to be pragmatic (cf. Yalcin, 2012). The types of obligation have proliferated, as variations on the 'ought'-operator – e.g., what one ought to do relative to a time and one's states of information, by contrast to what one ought to do relative to the facts – have been codified by differences in the array of intensional parameters relative to which the operator receives a semantic value (cf. Yalcin, op. cit.; Cariani, 2013; Dowell, 2013; et al).

This essay aims to provide a theory of the structural content of the types of intention via a similar modal analysis; to explain the role of intention in

---

[1]Deontic logic dates from at least as early as the fourteenth century, in the writings of Ockham, Holcot, and Rosetus. See Knuuttila (1981) for further discussion.

practical reasoning; and to answer thereby what I will call the unification problem: i.e., the inquiry into how the various types of intention comprise a unified mental state.[2] The general significance of the present contribution is that it will provide some foundational structure to the topic, where the previous lack thereof has served only to exacerbate its intransigence.[3] I will argue that – similarly to the case of deontic judgment – the foregoing types of intention can be countenanced as modal operators. The defining contours of the contents of the states may thus be targeted via their intensional-semantic profile. The types of intention on which I will focus include (i) the notion of 'intention-in-action', as evinced by cases in which agents act intentionally; (ii) the notion of 'intention-with-which', where an agent's intentions figure as an explanation of their actions; and (iii) the notion of 'intention-for-the-future', as evinced by an agent's plans to pursue a course of action at a future time.

I will argue that the unification problem has at least two, consistent solutions. The first manner in which the operations of intention are unified is that they are defined on a single space, whose points are states of information or epistemic possibilities. I argue, then, that the significance of examining how the state of intention interacts with practical reasoning is that it provides a second means by which to account for the unity of intention's types. Although each type of intention has a unique formal clause codifying its structural content, the notions of 'intention-in-action', 'intention-with-which', and 'intention-for-the-future' are nevertheless unified, because each is directed toward the property of expected utility. Thus, acting intentionally, acting because of an intention, and intending to pursue a course of action at a future time, are mental states whose unification consists in that each type aims toward the satisfaction of the value of an outcome – the value of

---

[2]The unification problem is first examined in Anscombe (1963), and has been pursued in contemporary research by, inter alia, Bratman (1984) and Setiya (2014).

[3]Compare the aims and methods pursued in the research projects of Fine (1981) and Williamson (2014): 'The relevance of the undertaking [. . .] consists mainly in the general advantages that accrue from formalizing an intuitive theory. First of all, one thereby obtains a clearer view of its primitive notions and truths. This is no small thing in a subject [. . .] that is so conspicuously lacking in proper foundations' (Fine, op. cit.:); 'The aim is to gain insight into a phenomenon by studying how it works under simplified, rigorously described conditions that enable us to apply mathematical or quasi-mathematical reasoning that we cannot apply directly to the phenomenon as it occurs in the wild, with all its intractable complexity. We can then cautiously transfer our insight about the idealized model back to the phenomenon in the wild' (Williamson, op. cit.).

which is the product of a partial belief conditional on one's acts by the utility thereof. The dissociation between an agent's intention to pursue an action and the causal relevance of the action's outcome adduces in favor of the characterization of expected utility in the setting of evidential, rather than causal, decision theory. The proposal that the content of intention is expected utility has, furthermore, the virtue of generalizing, in order to explain the nature of the intentions of non-human organisms. The contents of non-human organisms' intentions can here be understood as the value intended by their actions, as sensitive to both their prediction that the outcome will occur and the utility of its occurrence. Finally, because the aim of intention is expected utility, a precise account can be provided of how intention relates to the notions of belief and desire, while yet retaining its status as a unique mental state.

In Sections 2-3, I delineate the intensional-semantic profiles of the types of intention, and provide a precise account of how the types of intention are unified in virtue of both their operations in a single epistemic modal space and their role in practical reasoning, i.e., evidential decision theory. I endeavor to provide reasons adducing against the proposal that the types of intention are reducible to the mental states of belief and desire, where the former state is codified by subjective probability measures and the latter is codified by a utility function. Section 4 provides concluding remarks.

16.1   The Modes of Intention

The epistemic modal space of an agent can be defined via a frame, comprised of a set of points, and a relation of accessibility thereon (cf. Kripke, 1963; Blackburn et al, 2001). The points in the frame are here interpreted as an agent's states of information, while the relation of accessibility can receive various interpretations. A state of information is possible, just if there is at least one point relative to which it is true, if and only if it is not necessary for the formula to be false. One of the states of information is necessary, just if it is true everywhere, i.e. relative to all the other points in the space, if and only if it is impossible for it to be false. The distinctly epistemic interpretation of possibility comes in at least two guises, defined as the dual of epistemic necessity ('$\diamond\phi$' iff '$\neg\square\neg\phi$'): The truth of a formula is epistemically possible, just if the formula is believed by an agent, or is conceivable to the agent. The epistemic interpretation of necessity can itself come in at least two guises:

The truth of a formula is epistemically invariant or necessary, just if the truth of the formula is known by an agent, or if it is inconceivable for the formaula to be false (dually, epistemically necessary), and is thus in one sense apriori.

When an agent intends to ϕ, their intention may fall into three distinct types. One type of intention concerns the intentional pursuit, by the agent, of a course of action. A second type of intention can be witnessed, when the agent cites an intention as an explanation of her pursuit of a course of action. Finally, a third type of intention can be witnessed, when the agent intends to pursue a course of action at a future time.

16.1.1 Intention-in-Action

If the agent acts intentionally, then her intention can be understood as an operation relative to her states of information. The agent acts intentionally, just if there is a world and a unique array of intensional parameters relative to which her intention is realized and receives a positive semantic value. The array of intensional parameters is two-dimensional, because the value of intending to ϕ relative to one of the parameters will constrain the value of intending to ϕ relative to the subsequent parameters. Thus, we can say that an agent intends to ϕ, if and only if she acts intentionally, only if there is both a world and array of intensional parameters, relative to which her intention is realized, i.e. receives a positive value. The intensional parameters include a context comprised of a time and location, and a pair of indices on which spaces of the agent's acts and of the outcomes of her actions are built. So, the agent's intention-in-action receives a positive semantic value only if there is at least one world in her epistemic modal space at which – relative to the context of a particular time and location, which constrains the admissibility of the actions as defined at a first index, and which subsequently constrains the outcome thereof as defined at a second index – the intention is realized.

- $[\![\text{Intenton-in-Action}(\phi)]\!]_w = 1$ only if $\exists w'[\![\phi]\!]^{w',c(=t,l),a,o} = 1$

16.1.2 Intention-with-which

If the agent refers to an intention, in order to explain her pursuit of a course of action, then her intention can similarly be understood as an operation relative to her states of information. In this case, the agent intends to ϕ, just if there is a pair of formulas defined at points in her epistemic modal space,

where one of the states is realized because it holds in virtue the other state being realized. Informally, the foregoing explanation can be referred to as the intention-with-which she acts. Thus, we can say that an agent intends to ϕ, if and only if her intention is an explanation for her action, only if she acts in pursuit of ψ because she intends to ϕ. In order to capture the notion of one formula holding in virtue, or because, of a distinct formula, we define grounding operators on the agent's epistemic modal space. Thus, the agent intends to ϕ because, there is an intention in virtue of which her action, ψ so as to realize ϕ, receives a positive value.

- $[\![\text{Intention-with-which}(\phi)]\!]_w = 1$ only if $\exists w'[[\![\psi]\!]^{w'} = 1 \wedge [\![\text{G}(\phi,\psi)]\!] = 1]$,

where $\text{G}(x,y)$ is a grounding operator encoding the explanatory connection between ϕ and ψ. Following Fine (2012a,b), the grounding operator can have the following properties: The grounding operator is *weak* if and only if it induces reflexive grounding. The operator is *strict* if and only if it is not weak. The operator is *full* if and only if the intention to ϕ uniquely provides the explanatory ground for the action, ψ. The operator is *part* if and only if the intention to ϕ - along with other reasons for action - provide the explanatory ground for the action, ψ. Combinations of the foregoing explanatory operators may also obtain: $x < y$ iff ϕ is a *strict full* ground for ψ; $x \leq y$ iff ϕ is a *weak full* ground for ψ; $x \prec y$ iff ϕ is a *strict part* ground for ψ; $x \preceq y$ iff ϕ is a *weak part* ground for ψ; $x \preceq y \wedge \neg(y \preceq x)$ iff ϕ is a *strict partial* ground for ψ; $x \prec^* y$ iff $x_1, ..., x_n \leq y$, iff ϕ is a *partial strict* ground for ψ; $x \prec' z$ iff $[\phi \prec^* \psi \wedge \psi \preceq \mu]$ iff ϕ is a *part strict* ground for some further action, μ.

16.1.3   Intention-for-the-Future

Finally, an agent can intend to ϕ, because she intends to pursue a course of action at a future time. In this case, the intensional-semantic profile which records the parameters relative to which her intention receives a positive semantic value converges with a future-directed modal operator to the effect that the agent *will* ϕ. Thus, an agent realizes an intention-for-the-future only if there is a possible world and a future time, relative to which the possibility that ϕ is realized can be defined. Thus:

- $[\![\text{Intention-for-the-future}(\phi)]\!]_w = 1$ only if $\exists w'\forall t\exists t'[t< t' \wedge [\![\phi]\!]^{w',t'} = 1]$.[4]

This section has endeavored to accomplish two aims. The first was to provide a precise delineation of the structural content of, and therefore the distinctions between, the types of intention. Intention was shown to be a modal mental state, whose operations have a unique intensional profile, and whose values are defined relative to an agent's space of states of information. The second aim was to secure one of the means by which the unity of the distinct types of intention can be witnessed. Despite that each of the types of intention has a unique structural content, the contents of those types are each defined in a single, encompassing space; i.e, relative to the agent's space of epistemic possibilities.

16.2   Intention in Decision Theory

In Section 2, I suggested that intention is a unified, modal mental state, the contents of which are defined relative to an agent's states of information. This section examines the proposal that intentions have a dual profile (cf. Bratman, op. cit.), because intentions figure constitutively in practical reasoning. I argue that, because expected utility theories are the only axiomatized theories of practical reasoning, an account must be provided of the role that intention plays therein. The account will illuminate a precise relationship – which I argue is not identity – between the types of intention and the mental states of belief and desire. The account will furthur serve to provide a second explanation for the unity of intention's types, given the uniform role that the types of intention play in decision theory.

A model of decision theory can be understood as a tuple $\langle A,O,K,V\rangle$, where A is a set of acts; O is a set of outcomes; K encodes a set of counterfactual conditionals, where an act from A figures in the antecedent of the conditional and O figures in the conditional's consequent; and V is a function assigning a real number to each outcome. The real number is a representation of the value of the outcome. The expected value of the outcome is calculated

---

[4]See Rao and Georgeff (1991), for the suggestion that operators in a multi-modal logic can model the notion of goal-oriented intention. The foregoing intensional semantics is consistent with the logic that they proffer.

as the product of (i) the subjective probability – i.e., the agent's partial belief or credence – that the outcome will occur, as conditional on her act, and (ii) the value or utility which she assigns to the outcome's occurance. The agent can prefer one assignment of values to the outcome's occurrance over another. (Which preference axioms ought to be adopted is a contentious issue, and will not here be examined. Cf. von Neumann and Morgenstern, 1944; Savage, 1954; Jeffrey, 1983; and Joyce, 1999.) In evidential decision theory, the expected utility of an outcome is calculated as the product of the agent's credence conditional on her action, by the utility of the outcome. In causal decision theory, the expected utility of an outcome is calculated as the product of the agent's credence, conditional on both her action and the causal efficacy thereof, by the utility of the outcome. Expected utility can further be augmented by a risk-weighting function: If the agent's expected utility diminishes with the order of the bets she might pursue – such that expected utility is sensitive to the agent's propensity to take risks relative to the total ordering of the gambles – then she might have a preference for a sure-gain of .5 units of value, rather than prefer a bet with a 50 percent chance of winning either 0 or 1 units of value (cf. Buchak, 2014).

If intention plays a constitutive role in practical reasoning, and decision theories provide the most tractable models thereof, then what is the role of intention in decision theory? The parameters in the axiomatizations of decision theory encode variables for credences, actions, outcomes, assignments of utility, background states of information pertinent to the causal relevance of actions on outcomes, and the agent's preferences. Expected utility is derived, as noted, by the interaction between an agent's credences, actions, and utility assignments. Which, then, of these parameters do an agent's intentions concern?

There are dissociations between intention and belief and between intention and desire. An agent can have a partial belief that the sun will rise, without intending to pursue any course of action. Conversely, an agent can intend to pursue a course of action, yet appreciate that there are, unfortunately, reasons for her to disbelieve that the act will obtain. An agent can desire that the sun rises, without the intention to entrain the sun's rising as consequence. Conversely, a vegetarian can intend to consume meat, if it is the only available source of protein and they are in dire need thereof, while yet desire a distinct and orthogonal outcome.

There are dissociations between intention and preference. An agent can prefer the sun's rising to the prevalence in her life of unprovoked antagonists,

246

without either acting intentionally, possessing an intention as an explanation for some course of action, or intending to pursue any particular course of action in the future. Conversely, whether an agent's intention to pursue an action mandates a preference for the value of the outcome of that action will depend on one's preference axioms. One such axiom might be maximin, according to which the best of the worst outcomes among a set of options should be preferred, while a distinct rule might be maximax, according to which one ought to prefer and pursue the maximally vaulable outcome among a set of options. Thus, intending to ϕ is not sufficient for determining whether ϕ ought to be preferred.

There are, finally, dissociations between intention and action. One might intend to calculate the value of a formula, yet not be able so to act, because their attention might be allocated elsewhere.

Acting intentionally, intending to pursue a course of action in the future, and citing an intention as an explanation for one's course of action are each, however, in some way related to the value of a course of action. When an agent acts intentionally, she acts in such a way so as to obtain an outcome that she values. When an agent pursues a course of action, and refers to her intention so to act as the explanation for that action, the intention explains the value, for the agent, in which the action and its outcomes are supposed to consist. Finally, when an agent intends to pursue a course of action in the future, her intention is similarly guided by the value of the outcome that her action will hopefully entrain. The value of the outcome will not be her bare assessment of the utility of the outcome, because – in the setting of decision theory – utility functions codify desires, such that her intention would thereby be elided with her desire for the outcome.

Because the types of intention are all directed toward the value of an outcome of a course of action – while being irreducible to, because dissociable from, the states of belief and desire – the remaining and most suitable candidate for the role of the mental state of intention in decision theory is the aim of expected utility; i.e., the value of an outcome, as arising by the interaction between the agent's partial belief or expectation that the outcome will occur as conditional on her act, and the utility that she associates with the outcome's occurrence. Because of the dissociation between an agent's intention to pursue a course of action in the future and the action's occurrence – let alone the dissociation between the intention to act in future, and the causal efficacy of the action were it to obtain – the role of intention in practical reasoning appears to be more saliently witnessed in the setting of evidential

decision theory.

That the types of intention are each directed toward expected utility evinces how an agent's intentions can be sensitive to her beliefs and desires, without being reducible to them. Crucially, moreover, that the types of intention are each directed toward expected utility provides a second explanation of the way that the types of intention comprise a unified mental state.

Theoretical advantages accruing to the foregoing proposal include that it targets a foundational role for intention in decision theory. The proposal might be foundational, because it targets a basic role for intention in practical reasoning, which is consistent with the possible augmentation of the proposal with other approaches which assume a more cognitively demanding role for intention's aims. Such approaches include proposals to the effect (i) that the most fundamental type of intention is intention-with-which, such that intention's role as an explanation can be elided with its causal efficacy (cf. Anscombe, op. cit.; Davidson, 1963); (ii) that the content of intention is the diachronic satisfaction of self-knowledge (cf. Velleman, 1989); and (iii) that the role of intention in practical reasoning ought to be understood as an evaluative constraint on practical reasoning, as determined by the virtuous traits of an agent's character (cf. Setiya, 2007).

16.3 Concluding Remarks

I have argued that the unification problem for the types of intention can be solved in two, consistent ways. The types of intention can be modeled as modal operators, where the unity of the operations consists, in the first instance, in that their values are defined relative to a single, encompassing, epistemic modal space. The second manner by which the unity of intention's types can be witnessed is via intention's unique role in practical reasoning. I argued that each of the types of intention – i.e., intention-in-action, intention-as-explanation, and intention-for-the-future – has as its aim the value of an outcome of the agent's action, as derived by her partial beliefs and assignments of utility, and as codified by the value of expected utility in evidential decision theory. A precise account was thereby provided of the role of epistemic modality in the unification of the types of a unique, modal mental state, whose value figures constitutively in decision-making and practical reason.

Bibliography

Adams, R. 1979. Primitive Thisness and Primitive Identity. *Journal of Philosophy*, 76.

Angell, R. 1989. Deducibility, Entailment, and Analytic Entailment. In J. Norman and R. Sylvan (eds.), textitDirections in Relevant Logic. Kluwer Academic Publishers.

Anscombe, G.E.M. 1957/1963. *Intention*. Harvard University Press.

Anscombe, G.E.M. 1975. The First Person. In S. Guttenplan (ed.), *Mind and Language*. Clarendon Press.

Aristotle. 1987. *De Caelo*, tr. J.L. Stocks, revised J. Barnes (Revised Oxford Aristotle), text: D.J. Allan (Oxford Classical Texts, 1936). In J.L. Ackill (ed.), *A New Aristotle Reader*. Oxford University Press.

Aristotle. 1987. *De Interpretatione*, tr. J.L. Akrill (Clarendon Aristotle Series), text: L. Minio-Paluello (Oxford Classical Texts, 1956). In J.L. Ackill (ed.), *A New Aristotle Reader*. Oxford University Press.

Aristotle. 1987. *Physics*, tr. E. Hussey (Clarendon Aristotle Series, 1983), text: W.D. Ross (Oxford Classical Texts, 1950). In J.L. Ackill (ed.), *A New Aristotle Reader*. Oxford University Press.

Arntzenius, F. 2012. *Space, Time, and Stuff*. Oxford University Press.

Awodey, S. 2004. An Answer to Hellman's Question: 'Does Category Theory Provide a Framework for Mathematical Structuralism?'. *Philosophia Mathematica*, 12:3.

Awodey, S., L. Birkedal, and D. Scott. 2000. Local Realizability Toposes and a Modal Logic for Computability. Technical Report No. CMU-PHIL-99.

Awodey, S., and K. Kishida. 2007. Topology and Modality. http://repository.cmu.edu/philosophy/109.

Awodey, S., K. Kishida, and H.-C. Kotzsch. 2014. Topos Semantics for Higher-order Modal Logic. http://arxiv.org/abs/1403.0020v1.

Azzouni, J. 2004. Proof and Ontology in Euclidean Mathematics. In T. Kjeldsen, S. Pedersen, and L. Sonne-Hansen (eds.), *New Trends in the History and Philosophy of Mathematics*. University Press of Southern Denmark.

Azzouni, J. 2009. Why Do Informal Proofs Conform to Formal Norms? *Foundations of Science*, 14.

Azzouni, J. 2013. *Semantic Perception: How the Illusion of a Common Language Arises and Persists*. Oxford University Press.

Azzouni, J. Forthcoming. *Talking about Something: Metametaphysics, Quantifiers, and What There Is*.

Bacon, A. 2014. Representing Counterparts. *Australian Journal of Logic*, 11:2.

Bagaria, J., N. Castells, and P. Larson. 2006. An Ω-logic Primer. *Trends in Mathematics: Set Theory*. Birkhäuser Verlag.

Balog, K. 1999. Conceivability, Possibility, and the Mind-Body Problem. *Philosophical Review*, 108:4.

Baltag, A. 2003. A Coalgebraic Semantics for Epistemic Programs. *Electronic Notes in Theoretical Computer Science*, 82:1.

Barcan, R. 1946. A Functional Calculus of First Order Based on Strict Implication. *Journal of Symbolic Logic*, 11.

Barcan, R. 1947. The Identity of Individuals in a Strict Functional Calculus of Second Order. *Journal of Symbolic Logic*, 12.

Barcan Marcus, R. 1966. Iterated Deontic Modalities. *Mind*, 75.

Barnes, E., and J.R.G. Williams. 2011. A Theory of Metaphysical Indeterminacy. In K. Bennett and D. Zimmerman (eds.), *Oxford Studies in Metaphysics, Volume 6*. Oxford University Press.

Barsalou, L. 1999. Perceptual Symbol Systems. *Behavioral and Brain Sciences*, 22.

Barwise, J., and S. Feferman (eds.). 1985. *Model-Theoretic Logics*. Springer-Verlag.

Bayne, T., and M. Montague (eds.). 2011. *Cognitive Phenomenology*. Oxford University Press.

Bealer, G. 1982. *Quality and Concept.* Oxford University Press.

Beall, JC. 2000. Fitch's Proof, Verificationism, and the Knower Paradox. *Australasian Journal of Philosophy*, 78.

Beall, JC, and J. Murzi. 2012. Two Flavors of Curry's Paradox. *Journal of Philosophy*, 120.

Beddor, B. Forthcoming. Relativism and Expressivism. In Kusch (ed.), *Routledge Handbook to Relativism.* Routledge.

Belot, G. 2011. *Geometric Possibility.* Oxford University Press.

Benacerraf, P. 1965. What Numbers Could Not Be. *Philosophical Review*, 74.

Benacerraf, P. 1973. Mathematical Truth. *Journal of Philosophy*, 70.

Bengson, J. 2015. The Intellectual Given. *Mind*, 124:495.

van Benthem, J. 1983. *Modal Logic and Classical Logic.* Bibliopolis.

van Benthem, J. 1984/2003. Correspondence Theory. In D. Gabbay and F. Guenthner (eds.), *Handbook of Philosophical Logic, 2nd Edition, Volume 3.* Kluwer.

van Benthem, J. 2009. Actions that Make us Know. In Salerno (2009).

van Benthem, J. 2010. *Modal Logic for Open Minds.* CSLI.

Bermúdez, J. 2009. Truth, Indefinite Extensibility, and Fitch's Paradox. In Salerno (2009).

Bernays, P. 1942. A System of Axiomatic Set Theory: Part IV. General Set Theory. *Journal of Symbolic Logic*, 7:4.

Bernoulli, J. 1713/2006. *The Art of Conjecturing*, tr. E.D. Sylla. Johns Hopkins University Press.

Besson, C. 2010. Propositions, Dispositions, and Logical Knowledge. In M. Bonelli and A. Longo (eds.), *Quid Est Veritas?.* Bibliopolis.

Bird, A. 2007. *Nature's Metaphysics.* Oxford University Press.

Blackburn, P., and J. van Benthem. 2007. Modal Logic: A Semantic Perspective. In Blackburn, van Benthem, and F. Wolter (eds.), *Handbook of Modal Logic.* Elsevier.

Blackburn, P., M. de Rijke, and Y. Venema. 2001. *Modal Logic.* Cambridge University Press.

Blackburn, S. 1984. *Spreading the Word.* Oxford University Press.

Blanchette, 2012. *Frege's Conception of Logic.* Oxford University Press.

Block, N. 1983. Mental Pictures and Cognitive Science. *Philosophical Review*, 93.

Block, N. 2006. Max Black's Objection to Mind-Body Identity. In D. Zimmerman (ed.), *Oxford Studies in Metaphysics, Volume 3.* Oxford University Press.

Block, N. 2007. *Consciousness, Function, and Representation.* MIT Press.

Block, N. 2007. Sexism, Racism, Ageism, and the Nature of Consciousness. In Block, *Consciousness, Function, and Representation.* MIT Press.

Block, N. 2010. Attention and Mental Paint. *Philosophical Issues*, 20:1.

Block, N. 2011. Response to Rosenthal and Weisberg. *Analysis*, 71:3.

Block, N. 2013. The Grain of Vision and the Grain of Attention. *Thought*, 1:2.

Block, N. 2015. The Puzzle of Perceptual Precision. In T. Metzinger and J. M. Windt (eds), *Open MIND*: 5(T). MIND Group.

Block, N. 2015. The Canberra Plan Neglects Ground. In T. Horgan, M. Sabates, and D. Sosa (eds.), *Qualia and Mental Causation in a Physical World: Themes from the Philosophy of Jaegwon Kim.* Cambridge University Press.

Block, N., and R. Stalnaker. 1999. Conceptual Analysis, Dualism, and the Explanatory Gap. *Philosophical Review*, 108:1.

Boghossian, P. 2003. Blind Reasoning. *Proceedings of the Aristotelian Society, Supplementary Volume*, 77.

Boh, I. 1993. *Epistemic Logic in the Later Middle Ages.* Routledge.

Bolzano, B. 1810/2004. *Contributions to a Better-grounded Presentation of Mathematics*, tr. S. Russ. In Russ, *The Mathematical Works of Bolzano.* Oxford University Press.

BonJour, L. 1998. *In Defense of Pure Reason.* Cambridge University Press.

Boole, G. 1854. *The Laws of Thought.* London: Walton and Maberly. Cambridge: MacMillan.

Boolos, G. 1971. The Iterative Conception of Set. *Journal of Philosophy*, 68.

Boolos, G. 1984. To Be is to Be the Value of a Variable (or to Be some Values of some Variables). *Journal of Philosophy*, 81.

Boolos, G. 1985. Nominalist Platonism. *Philosophical Review*, 94:3.

Boolos, G. 1987/1998. The Consistency of Frege's *Foundations of Arithmetic*. In J.J. Thomson (ed.), *On Being and Saying*. MIT Press. Reprinted in Boolos, *Logic, Logic, and Logic*, ed. R. Jeffrey. Harvard University Press.

Boolos, G. 1990/1998. The Standard of Equality of Numbers. In Boolos (ed.), *Meaning and Method*. Cambridge University Press. Reprinted in Boolos (1998).

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

Boolos, G. 1997. Is Hume's Principle Analytic? In R. Heck (ed.), *Language, Thought, and Logic*. Oxford University Press.

Bradwardine, T. c.1320/2010. *Insolubilia*, tr. and ed. S. Read. Dallas Medieval Texts and Translations, 10.

Brandom, R. 2008. *Between Saying and Doing*. Oxford University Press.

Brandom, R. 2014. *From Empiricism to Expressivism*. Harvard University Press.

Bratman, M. 1984. Two Faces of Intention. *Philosophical Review*, 93:3.

Briggs, R. 2012. Interventionist Counterfactuals. *Philosophical Studies*, 160:1.

Brighouse, C. 1999. Incongruent Counterparts and Modal Relationalism. *International Studies in the Philosophy of Science*, 13:1.

Bub, J. 2011. Quantum Probabilities: An Information-Theoretic Interpretation. In C. Beisbart and S. Hartmann (eds.), *Probabilities in Physics*. Oxford University Press.

Buchak, L. 2014. *Risk and Rationality*. Oxford University Press.

Burali-Forti, C. 1897/1967. A Question on Transfinite Numbers (tr. J. van Heijenoort). In van Heijenoort (ed.), *From Frege to Gödel*. Harvard University Press.

Burge, T. 2005. *Truth, Thought, Reason: Essays on Frege*. Oxford University Press.

Burge, T. 2010. *Origins of Objectivity*. Oxford University Press.

Burgess, J. 2004/2008. *E Pluribus Unum*: Plural Logic and Set Theory. *Philosophia Mathematica*, 12:3. Reprinted in Burgess (2008), *Mathematics, Models, and Modality*. Oxford University Press.

Burgess, J. 2014. Intuitions of Three Kinds in Gödel's Views on the Continuum. In J. Kennedy (ed.), *Interpreting Gödel*. Cambridge University Press.

Burgess, J., and G. Rosen. 1997. *A Subject with No Object*. Oxford University Press.

Buridan, J. 2001. *Summulae de Dialectica*, tr. G. Klima. Yale University Press.

Butterfield, J. 2004. Some Aspects of Modality in Analytical Mechanics. In M. Stöltzner and P. Weingartner (eds.), *Formal Teleology and Causation*. Mentis.

Byrne, A. 2009. Experience and Content. *Philosophical Quarterly*, 59.

Byrne, A., and D. Hilbert. 2007. Color Primitivism. *Erkenntnis*, 66.

Byrne, R. 2002. Mental Models and Counterfactual Thoughts about What Might Have Been. *Trends in Cognitive Sciences*, 6:10.

Byrne, R. 2005. *The Rational Imagination*. MIT Press.

Caie, M. 2011. Belief and Indeterminacy. *Philosophical Review*, 121:1.

Caie, M. 2013. Rational Probabilistic Incoherence. *Philosophical Review*, 122:4.

Camp, E. 2007. Thinking with Maps. *Philosophical Perspectives*, 21.

Cantor, G. 1883/1996. *Foundations of a General Theory of Manifolds*, tr. W.B. Ewald. In Ewald (ed.), *From Kant to Hilbert: A Source Book in the Foundations of Mathematics, Vol. II*. Oxford University Press.

Cantor, G. 1891/1996. 'On an Elementary Question in the Theory of Manifolds', tr. Ewald. In Ewald (op. cit.).

Cantor, G. 1895/2007. *Contributions to the Founding of the Theory of Transfinite Numbers*, tr. P. Jourdain. Cosimo.

Cantor, G. 1899/1967. 'Letter to Dedekind', tr.S. Bauer-Mengelberg and J. van Heijenoort. In van Heijenoort (ed.), *From Frege to Gödel*. Harvard University Press.

Cappelen, H. 2012. *Philosophy without Intuitions*. Oxford University Press.

Carandini, M., and D. Heeger. 2011. Normalization as a Canonical Neural Computation. *Nature Reviews Neuroscience*, 13.1.

Cariani, F. 2013. 'Ought' and Resolution Semantics. *Nous*, 47:3.

Cariani, F., M. Kaufmann, and S. Kaufmann. 2013. Deliberative Modality under Uncertainty. *Linguistics and Philosophy*, 36.

Carnap, R. 1937. *The Logical Syntax of Language*, tr. A. Smeaton. Routledge.

Carnap, R. 1945. On Inductive Logic. *Philosophy of Science*, 12:2.

Carnap, R. 1947. *Meaning and Necessity.* University of Chicago Press.

Carrasco, M., S. Ling, and S. Read. 2004. Attention alters Appearance. *Nature Neuroscience*, 7.3.

Carrasco, M. 2011. Visual Attention: The Past 25 Years. *Vision Research*, 51.

Carruthers, P. 1996. *Language, Thought and Consciousness.* Cambridge University Press.

Chalmers, D. 1996. *The Conscious Mind.* Oxford University Press.

Chalmers, D. 2002. Does Conceivability Entail Possibility? In T. Gendler and J. Hawthorne (eds.), *Conceivability and Possibility.* Oxford University Press.

Chalmers, D. 2004. Epistemic Two-dimensional Semantics. *Philosophical Studies*, 118.

Chalmers, D. 2006. The Foundations of Two-dimensional Semantics. In M. Garcia-Carpintero and J. Macia (eds.), *Two-dimensional Semantics.* Oxford University Press.

Chalmers, D. 2010. *The Character of Consciousness.* Oxford University Press.

Chalmers, D. 2011. The Nature of Epistemic Space. In A. Egan and B. Weatherson (eds.), *Epistemic Modality.* Oxford University Press.

Chalmers, D. 2012. *Constructing the World.* Oxford University Press.

Chalmers, D. 2012. The Varieties of Computation. *Journal of Cognitive Science*, 13.

Chalmers, D. 2013. The Contents of Consciousness. *Analysis*, 73:2.

Chalmers, D. 2014. Strong Necessities and the Mind-Body Problem. *Philosophical Studies*, 167:3.

Chalmers, D. 2014. Intensions and Indeterminacy. *Philosophy and Phenomenological Research*, 89:1.

Chalmers, D., and B. Rabern. 2014. Two-dimensional Semantics and the Nesting Problem. *Analysis*, 74:2.

Chang, R. 2013. Grounding Practical Normativity. *Philosophical Studies*, 164:1.

Chihara, C. 1990. *Constructibility and Mathematical Existence*. Oxford University Press.

Chihara, C. 1998. Tarski's Thesis and the Ontology of Mathematics. In M. Schirn (ed.), *Philosophy of Mathematics Today*. Oxford University Press.

Chihara, C. 2004. *A Structural Account of Mathematics*. Oxford University Press.

Chirimuuta, M. 2014. Minimal Models and Canonical Neural Computations. *Synthese*, 191:2.

Chudnoff, E. 2013. *Intuition*. Oxford University Press.

Church, A. 1936. An Unsolvable Problem of Elementary Number Theory. *American Journal of Mathematics*, 58:2.

Church, A. 1954. Intensional Isomorphism and the Identity of Belief. *Philosophical Studies*, 5:5.

Church, A. 1993. A Revised Formulation of the Logic of Sense and Denotation. *Noûs*, 27:2.

Church, A. 2009. Referee Reports on Fitch's 'A Definition of Value'. In Salerno (2009).

Clark, P. 2007. Frege, Neo-Logicism, and Applied Mathematics. In R. Cook (ed.), *The Arché Papers on the Mathematics of Abstraction*. Springer.

Clifford, W.K. 1877. The Ethics of Belief. *Contemporary Review*, 29.

Cook, R. 2014. *The Yablo Paradox*. Oxford University Press.

Cook, R. 2014. Possible Predicates and Actual Properties. *Synthese*, DOI 10.1007/s11229-014-0592-1.

Cook, R. 2016. Necessity, Necessitism, and Numbers. *Philosphical Forum*, 47: 3-4.

Cook, R., and P. Ebert. 2005. Abstraction and Identity. *Dialectica*, 59:2.

Correia, F., and B. Schnieder (eds.). 2012. *Metaphysical Grounding*. Cambridge University Press.

Cotnoir, A. 2009. Anti-symmetry and Non-extensional Mereology. *Philosophical Quarterly*, 60.

Cotnoir, A. 2014. Does Universalism Entail Extensionalism? *Nous*, DOI: 10.1111/nous.12063.

Cotnoir, A. 2015. Abelian Mereology. *Logic and Logical Philosophy*, DOI: 10.12775/LLP.2015.006.

Cotnoir, A. ms. Are Ordinary Objects Abstracta?

Daston, L. 1988. *Classical Probability in the Enlightenment*. Princeton University Press.

Daston, L. 1994. How Probabilities came to be Objective and Subjective. *Historia Mathematica*, 21.

Davidson, D. 1963. Actions, Reasons, and Causes. *Journal of Philosophy*, 23.

De Brigard, F., D.R. Addis, J.H. Ford, D.L. Schachter, and D.S. Giovanello. 2013. Remembering What Could Have Happened: Neural Correlates of Episodic Counterfactual Thinking. *Neuropsychologia*, 51.12.

Dean, W. 2015. Computational Complexity Theory. *Stanford Encyclopedia of Philosophy*.

Dean, W. 2016. Algorithms and the Mathematical Foundations of Computer Science. In L. Horsten and P. Welch (eds.), *Gödel's Disjunction*. Oxford University Press.

Dedekind, R. 1872/1996. 'Continuity and Irrational Numbers', tr. Ewald. In Ewald (1996).

Dedekind, R. 1888/1996. *Was sind und was sollen die Zahlen?*, tr. Ewald. In Ewald (1996).

de Finetti, B. 1937/1964. *Foresight: Its Logical Laws, Its Subjective Sources*, tr. H. Kyburg. In Kyburg and H. Smokler (eds.), *Studies in Subjective Probability*. Wiley.

Detlefsen, M. 1985. Wright on the Non-mechanizability of Intuitionist Reasoning. *Philosophia Mathematica*, 3:3.

Deutsch, D. 2013. Constructor Theory. *Synthese*, 190.

Dever, J. ms. Quantity without Quantities.

Dorr, C. 2003. Vagueness without Ignorance. *Philosophical Perspectives*, 17.

Dorr, C. ms. Finding Ordinary Objects in the World of Quantum Mechanics.

Dorr, C., and J. Goodman. ms. Diamonds are Forever.

Dowell, J.L. 2013. Flexible Contextualism about Deontic Modals. *Inquiry*, 56.

Dretske, F. 1995. *Naturalizing the Mind*. MIT Press.

Dummett, M. 1959. Truth. *Proceedings of the Aristotelian Society, Supplementary Volume*, 59.

Dummett, M. 1963/1978. The Philosophical Significance of Gödel's Theorem. In Dummett (1978), *Truth and Other Enigmas*. Harvard University Press.

Dummett, M. 1977/2000. *Elements of Intuitionism*. Oxford University Press.

Dummett, M. 1978. *Truth and Other Enigmas*. Harvard University Press.

Dummett, M. 1991. *The Logical Basis of Metaphysics*. Harvard University Press.

Dummett, M. 1991. *Frege: Philosophy of Mathematics*. Duckworth.

Dummett, M. 1996. What is Mathematics about? In Dummett, *The Seas of Language*. Oxford University Press.

Eagle, A. 2008. Mathematics and Conceptual Analysis. *Synthese*, 161:1.

Easwaran, K. 2014. Decision Theory without Representation Theorems. *Philosophers' Imprint*, 14:27.

Easwaran, K., and B. Fitelson. 2015. Accuracy, Coherence, and Evidence. In T. Gendler and J. Hawthorne (eds.), *Oxford Studies in Epistemology, Volume 5*. Oxford University Press.

Edgington, D. 2010. Possible Knowledge of Unknown Truth. *Synthese*, 173.

Edgington, D. 2014. Two Kinds of Possibility. *Proceedings of the Aristotelian Society, Supplementary Volumes*, Volume 78.

Eklund, M. 2006. Neo-Fregean Ontology. *Philosophical Perspectives*, 20.

Eklund, M. 2016. Hale and Wright on the Metaontology of Neo-Fregeanism. In P. Ebert and M. Rossberg (eds.), *Abstractionism: Essays in Philosophy of Mathematics*. Oxford University Press.

Egan, A., and B. Weatherson. 2011. Introduction, in *Epistemic Modality*. Oxford University Press.

Egan, F. 2010. Computational Models. *Studies in History and Philosophy of Science*, Part A, 41:3.

Enoch, D., and J. Schechter. 2008. How Are Basic Belief-forming Methods Justified? *Philosophy and Phenomenological Research*, 76:3.

Etchemendy, J. 1990/1999. *The Concept of Logical Consequence*. CSLI Publications.

Evans, G. 1979. Reference and Contingency. *The Monist*, 62:2.

Evans, G. 1982. *The Varieties of Reference*, ed. J. McDowell. Oxford University Press.

Evans, G. 1985. *Collected Papers*. Oxford University Press.

Evans, G. 2004. Comment on 'Two Notions of Necessity'. *Philosophical Studies*, 118:1/2.

Ewald, W.B. (ed.). 1996. *From Kant to Hilbert: A Source Book in the Foundations of Mathematics, Vol. II*. Oxford University Press.

Fagin, R., J. Halpern, Y. Moses, and M. Vardi. 1995. *Reasoning about Knowledge*. MIT Press.

Fara, D.G. 2000. Shifting Sands. *Philosophical Topics*, 28. Originally published under the name 'Delia Graff'.

Fara, D.G. 2001. Phenomenal Continua and the Sorites. *Mind*, 110. Originally published under the name 'Delia Graff'.

Fara, D.G. 2008. Profiling Interest Relativity. *Analysis*, 68:4.

Fara, D.G. 2008. Relative-Sameness Counterpart Theory. *Review of Symbolic Logic*, 1:2.

Feferman, S. 1984. Toward Useful Type-free Theories, I. *Journal of Symbolic Logic*, 49:1.

Feferman, S. 1999. Logic, Logics, and Logicism. *Notre Dame Journal of Formal Logic*, 40:1.

Field, H. 1980/2016. *Science without Numbers*, 2nd ed. Oxford University Press.

Field, H. 1989. *Realism, Mathematics, and Modality*. Basil Blackwell.

Field, H. 2001. *Truth and the Absence of Fact*. Oxford University Press.

Field, H. 2004. The Consistency of the Naive Theory of Properties. *Philosophical Quarterly*, 54:214.

Field, H. 2008. *Saving Truth from Paradox*. Oxford University Press.

Field, H. 2015. What is Logical Validity? In C. Caret and O. Hjortland (eds.), *Foundations of Logical Consequence*. Oxford University Press.

Fine, K. 1974. Logics containing K4: Part I. *Journal of Symbolic Logic*, 39:1.

Fine, K. 1975. Some Connections between Elementary and Modal Logic. In S. Kanger (ed.), *Proceedings of the Third Scandinavian Logic Symposium*. North Holland.

Fine, K. 1977. Properties, Propositions and Sets. *Journal of Philosophical Logic*, 6.

Fine, K. 1981. First-order Modal Theories I – Sets. *Nous*, 15:2.

Fine, K. 1985. *Reasoning with Arbitrary Objects*. Blackwell Publishing.

Fine, K. 1994. Essence and Modality. *Philosophical Perspectives*, 8.

Fine, K. 1985. The Logic of Essence. *Journal of Philosophical Logic*, 24:3.

Fine, K. 2001. The Question of Realism. *Philosophers' Imprint*, 1:1.

Fine, K. 2002. *The Limits of Abstraction*. Oxford University Press.

Fine, K. 2005. *Modality and Tense*. Oxford University Press.

Fine, K. 2005. Our Knowledge of Mathematical Objects. In T. Gendler and J. Hawthorne (eds.), *Oxford Studies in Epistemology, Volume 1*. Oxford University Press.

Fine, K. 2005. Class and Membership. *Journal of Philosophy*, 102:11.

Fine, K. 2006. Relatively Unrestricted Quantification. In A. Rayo and G. Uzquiano (eds.), *Absolute Generality*. Oxford University Press.

Fine, K. 2007. Response to Weir. *Dialectica*, 61:1.

Fine, K. 2009. *Semantic Relationalism*. Blackwell.

Fine, K. 2010. Semantic Necessity. In B. Hale and A. Hoffman (eds.), *Modality*. Oxford University Press.

Fine, K. 2012. A Difficulty for the Possible Worlds Analysis of Counterfactuals. *Synthese*, DOI 10.1007/s11229-012-0094-y.

Fine, K. 2012. Counterfactuals without Possible Worlds. *Journal of Philosophy* 109:3.

Fine, K. 2012. The Pure Logic of Ground. *Review of Symbolic Logic*, 5:1.

Fine, K. 2012. Guide to Ground. In F. Correia and B. Schnieder (eds.), *Metaphysical Grounding*. Cambridge University Press.

Fine, K. 2013. Truth-maker Semantics for Intuitionistic Logic. *Journal of Philosophical Logic*, 43.

Fine, K. 2015. Angellic Content. *Journal of Philosophical Logic*, DOI 10.1007/s10992-015-9371-9.

Fine, K. 2015. Identity Criteria and Ground. *Philosophical Studies*, DOI 10.1007/s11098-014-0440-7.

Fine, K. 2015. Unified Foundations for Essence and Ground. *Journal of the American Philosophical Association*, 1:2.

Fine, K. 2017a. A Theory of Truthmaker Content I: Conjunction, Disjunction, and Negation. *Journal of Philosophical Logic*, 46:6.

Fine, K. 2017b. A Theory of Truthmaker Content II: Subject-matter, Common Content, Remainder, and Ground. *Journal of Philosophical Logic*, 46:6.

von Fintel, K., and I. Heim. 2011. *Intensional Semantics*. MIT.

Fitch, F. 1963. A Logical Analysis of Some Value Concepts. *Journal of Symbolic Logic*, 28.

Fitelson, B. 2001. *Studies in Confirmation Theory*. Ph.D. Dissertation, University of Wisconsin - Madison.

Flagg, R. 1985. Epistemic Set Theory is a Conservative Extension of Intuitionistic Set Theory. *Journal of Symbolic Logic*, 50:4.

Fodor, J. 1975. *The Language of Thought*. Harvard University Press.

Fodor, J. 1980. Methodological Solipsism considered as a Research Strategy in Cognitive Psychology. *Behavioral and Brain Sciences*, 3.

Fodor, J. 2000. *The Mind Doesn't Work that Way*. MIT Press.

Fodor, J. 2008. *LOT 2.* Oxford University Press.

Frege, G. 1884/1980. *The Foundations of Arithmetic*, 2nd ed., tr. J.L. Austin. Northwestern University Press.

Frege, G. 1891. 'Function and Concept', tr. P. Geach. In Frege (1997), *The Frege Reader*, ed. M. Beaney. Blackwell Publishing.

Frege, G. 1892. 'On *Sinn* and *Bedeutung*', tr. M. Black. In Frege (1997).

Frege, G. 1893/2013. *Basic Laws of Arithmetic, Vol. I-II*, tr. and ed. P. Ebert, M. Rossberg, C. Wright, and R. Cook. Oxford University Press.

Frege, G. 1897. 'Logic', tr. P. Long and R. White. In Frege (1997).

Friedman, H. 1975/ms. The Analysis of Mathematical Texts, and Their Calibration in Terms of Intrinsic Strength I.

Friedman, M. 2001. *Dynamics of Reason.* CSLI Publications.

Frith, C., S.-J. Blakemore, and D. Wolport. 2000. Abnormalities in the Awareness and Control of Action. *Philosophical Transactions of the Royal Society of London B*, 355.

Fritz, P. 2016. First-order Modal Logic in the Necessary Framework of Objects. *Canadian Journal of Philosophy*, 46:4-5.

Fritz, P., and J. Goodman. 2017. Counting Incompossibles. *Mind*, 126:504.

Galles, D., and J. Pearl. 1998. An Axiomatic Characterization of Causal Counterfactuals. *Foundations of Science*, 3:1.

Gendler, T. 2010. *Intuition, Imagination, and Philosophical Methodology.* Oxford University Press.

Gendler, T., and J. Hawthorne (eds.). 2002. *Conceivability and Possibility.* Oxford University Press.

Giaquinto, M. 2008. Visualizing in Mathematics. In P. Mancosu (ed.), *The Philosophy of Mathematical Practice.* Oxford University Press.

Gibbard, A. 2012. *Meaning and Normativity.* Oxford University Press.

Gödel, K. 1931. On Formally Undecidable Propositions of *Principia Mathematica* and Related Systems I. In Gödel (1986), *Collected Works, Volume I*, eds. S. Feferman, J. Dawson, S. Kleene, G. Moore, R. Solovay, and J. van Heijenoort. Oxford University Press.

Gödel, K. 1933. An Interpretation of the Intuitionistic Propositional Calculus. In Gödel (1986).

Gödel, K. 1938. Lecture at Zilsel's. In Gödel (1995), *Collected Works, Volume III*, eds. S. Feferman, J. Dawson, W. Goldfarb, C. Parsons, and R. Solovay. Oxford University Press.

Gödel, K. 1944. Russell's Mathematical Logic. In Gödel (1990), *Collected Works, Volume II*, eds. S. Feferman, J. Dawson, S. Kleene, G. Moore, R. Solovay, and J. van Heijenoort. Oxford University Press.

Gödel, K. 1946. Remarks before the Princeton Bicentennial Conference on Problems in Mathematics. In Gödel (1990).

Gödel, K. 1951. Some Basic Theorems on the Foundations of Mathematics and their Implications. In Gödel (1995).

Gödel, K. 1953,III-V. Is Mathematics Syntax of Language? In Gödel (1995).

Gödel, K. 1961. The Modern Development of the Foundations of Mathematics in the Light of Philosophy. In Gödel (1995).

Gödel, K. 1947/1964. What is Cantor's Continuum Problem? In Gödel (1990).

Goldblatt, R. 1987. *Logics of Time and Computation.* CSLI.

Goldblatt, R. 1993. *Mathematics of Modality.* CSLI Publications.

Goldblatt, R. 2003. Questions of Canonicity. In V. Hendricks and J. Malinowski (eds.), *Trends in Logic – 50 Years of Studia Logica.* Kluwer.

Goldblatt, R. forthcoming. Fine's Theorem on First-order Complete Modal Logics. In M. Dumitru (ed.), *Metaphysics, Meaning and Modality: Themes from Kit Fine.* Oxford University Press.

Gómez-Torrente, M. 1996. Tarski on Logical Consequence. *Notre Dame Journal of Formal Logic*, 37:1.

Gómez-Torrente, M. 2002. Vagueness and Margin for Error Principles. *Philosophy and Phenomenological Research*, 64:1.

Goodman, N.D. 1985. A Genuinely Intensional Set Theory. In Shapiro (ed.), *Intensional Mathematics.* Elsevier Science Publishers.

Goodman, N.D. 1990. Topological Models of Epistemic Set Theory. *Annals of Pure and Applied Mathematics*, 46.

Gottlieb, J., and P. Balan. 2010. Attention as a Decision in Information Space. doi:10.1016/j.tics.2010.03.001

Gurevich, Y. 1999. The Sequential ASM Thesis. *Bulletin of the European Association for Theoretical Computer Science*, 67.

Halbach, V., and A. Visser. 2014. Self-reference in Arithmetic I. *Review of Symbolic Logic*, 7:4.

Hale, B. 1987. *Abstract Objects*. Basil Blackwell.

Hale, B. 2000. Reals by Abstraction. *Philosophia Mathematica*, 3:8. Reprinted in Hale and Wright (2001).

Hale, B. 2000. Abstraction and Set Theory. *Notre Dame Journal of Formal Logic*, 41:4.

Hale, B. 2013. *Necessary Beings*. Oxford University Press.

Hale, B. 2013. Properties and the Interpretation of Second-order Logic. *Philosophia Mathematica*, 3:21.

Hale, B., and C. Wright. 2001. *The Reason's Proper Study*. Oxford University Press.

Hale, B., and C. Wright. 2002. Benaceraff's Dilemma Revisited. *European Journal of Philosophy*, 10:1.

Hale, B. and C. Wright. 2005. Logicism in the Twenty-first Century. In S. Shapiro (ed.), *The Oxford Handbook of Philosophy of Mathematics and Logic*. Oxford University Press.

Hale, B., and C. Wright. 2009. The Metaontology of Abstraction. In D. Chalmers, D. Manley, and R. Wasserman (eds.), *Metametaphysics*. Oxford University Press.

Hamkins, J. 2012. The Set-theoretic Multiverse. *Review of Symbolic Logic*, 5:3.

Hamkins, J., and B. Löwe. 2008. The Modal Logic of Forcing. *Transactions of the American Mathematical Society*, 360:4.

Hamkins, J., and B. Löwe. 2013. Moving Up and Down in the Generic Multiverse. In K. Lodaya (ed.), *Logic and Its Applications*. Springer.

Hardin, C.L. 1988. *Color for Philosophers*. Hackett Publishing.

Hardy, G., and E.M. Wright. 1979. *An Introduction to the Theory of Numbers*, 5th ed. Oxford University Press.

Haugeland, J. 1978. The Nature and Plausibility of Cognitivism. *Behavioral and Brain Sciences*, 2.

Hawke, P., and S. Steinert-Threlkeld. 2020. Semantic Expressivism for Epistemic Modals. *Linguistics and Philosophy*, https://doi.org/10.1007/s10988-020-09295-7.

Hawley, K. 2007. Neo-Fregeanism and Quantifier Variance. *Proceedings of the Aristotelian Society Supplementary Volume*, 81.

Hawthorne, J. 2005. Vagueness and the Mind of God. *Philosophical Studies*, 122.

Hawthorne, J. 2010. A Metaphysician Looks at the Everett Interpretation. In S. Saunders, J. Barrett, A. Kent, and D. Wallace (eds.), *Many Worlds? Everett, Quantum Theory, and Reality*. Oxford University Press.

Hazen, A. 1985. Review of Crispin Wright's *Frege's Conception of Numbers as Objects*. *Australasian Journal of Philosophy*, 63:2.

Heck, R. 1992. On the Consistency of Second-order Contextual Definitions. *Nous*, 26.

Heck, R. 2011. *Frege's Theorem*. Oxford University Press.

Heim, I. 1992. Presupposition Projection and the Semantics of Attitude Verbs. *Journal of Semantics*, 9.

Hellman, G. 1990. Toward a Modal-Structural Interpretation of Set Theory. *Synthese*, 84.

Hellman, G. 1993. *Mathematics without Numbers*. Oxford University Press.

Helmholtz, H. von. 1878/1977. The Facts in Perception. In Helmholtz, *Epistemological Writings*, tr. M. Lowe, eds. R. Cohen and Y. Elkana. D. Reidel Publishing.

Henkin, L. 1952. A Problem concerning Provability. *Journal of Symbolic Logic*, 17:2.

Henkin, L., J.D. Monk, and A. Tarski. 1971. *Cylindric Algebras*, Part I. North-Holland.

Hicks, M.T., and J. Schaffer. 2015. Derivative Properties in Fundamental Laws. *British Journal for the Philosophy of Science*, doi:10.1093/bjps/axv039.

Hilbert, D. Letter from Hilbert to Frege, 29.12.1899. In G. Frege, *Philosophical and Mathematical Correspondence*, tr. H. Kaal, ed., G. Gabriel, H. Hermes, F. Kambartel, C. Thiel, and A. Veraart. Basil Blackwell.

Hilbert, D. 1923/1996. 'The Logical Foundations of Mathematics', tr. Ewald. In Ewald (1996).

Hill, C. 1997. Imaginability, Conceivability, Possibility, and the Mind-Body Problem. *Philosophical Studies*, 87:1.

Hill, C. 2006. Modality, Modal Epistemology, and the Metaphysics of Consciousness. In S. Nichols (ed.), *The Architecture of the Imagination*. Oxford University Press.

Hilpinen, R. (ed.). 1981. *New Studies in Deontic Logic*. D. Reidel Publishing.

Hintikka, J. 1962. *Knowledge and Belief*. Cornell University Press.

Hjortland, O. 2016. Anti-exceptionalism about Logic. *Synthese*, DOI 10.1007/s11098-016-0701-8.

Hoare, C. 1969. An Axiomatic Basis for Computer Programming. *Communications of the Association for Computing Machinery*, 12.

Hodes, H. 1984,a. Logicism and the Ontological Commitments of Arithmetic. *Journal of Philosophy*, 81:3.

Hodes, H. 1984,b. On Modal Logics which Enrich First-Order S5. *Journal of Philosophical Logic*, 13:4.

Hornsby, J. 1997. The Presidential Address: Truth: The Identity Theory. *Proceedings of the Aristotelian Society, New Series*, Vol. 97.

Horsten, L. 2010. Perceptual Indiscriminability and the Concept of a Color Shade. In R. Dietz and S. Moruzzi (eds.), *Cuts and Clouds*. Oxford University Press.

Horty, J. 2001. *Agency and Deontic Logic*. Oxford University Press.

Horty, J., and N. Belnap. 1995. The Deliberative Stit. *Journal of Philosophical Logic*, 24:6.

Hughes, J. 2001. *A Study of Categories of Algebras and Coalgebras*. Ph.D. Dissertation, Carnegie Mellon University.

Hume, D. 1739-1740/2007. *A Treatise of Human Nature*, eds. D.F. Norton and M.J. Norton. Oxford University Press.

Husserl, E. 1929/1999. *Formal and Transcendental Logic*, tr. D. Cairns. In D. Welton (ed.), *The Essential Husserl*. Indiana University Press.

Ichikawa, J., and B. Jarvis. 2013. *The Rules of Thought*. Oxford University Press.

van Inwagen, P. 1990. *Material Beings*. Cornell University Press.

Jackson, F. 1982. Epiphenomenal Qualia. *Philosophical Quarterly*, 32.

Jackson, F. 2011. Possibilities for Representation and Credence. In A. Egan and B. Weatherson (eds.), *Epistemic Modality*. Oxford University Press.

Jackson, F., K. Mason, and S. Stich. 2008. Folk Psychology and Tacit Theories. In D. Braddon-Mitchell and R. Nola (eds.), *Conceptual Analysis and Philosophical Naturalism*. MIT Press.

James, W. 1884. On Some Omissions of Introspective Psychology. *Mind*, 33.

James, W. 1896. The Will to Believe. *The New World*, 5.

Janin, D., and I. Walukiewicz. 1996. On the Expressive Completeness of the Propositional μ-Calculus with respect to Monadic Second-order Logic. In U. Montanari and V. Sasone (eds.), *Proceedings CONCUR, '96*. Springer.

Jech, T. 2003. *Set Theory*, 3rd Millennium ed. Springer.

Jeffrey, R. 1965/1990. *The Logic of Decision*. University of Chicago Press.

Jennings, C.D. 2012. The Subject of Attention. *Synthese*, 189:3.

Jennings, C.D. 2014. Attention and Perceptual Organization. *Philosophical Studies*, DOI 10.1007/s11098-014-0348-2.

Jiang, Y., P. Costello, F. Fang, M. Huang, and S. He. 2006. A Gender and Sexual Orientation-dependent Spatial Attentional Effect of Invisible Images. *Proceedings of the National Academy of Sciences*, 103:45.

Johnson-Laird, P.N. 2004. The History of Mental Models. In K. Manktelow and M. Chung (eds.), *Psychology of Reasoning*. Psychology Press.

Joyce, J. 1998. A Non-pragmatic Vindication of Probabilism. *Philosophy of Science*, 65:4.

Joyce, J. 1999. *The Foundations of Causal Decision Theory*. Cambridge University Press.

Joyce, J. 2009. Accuracy and Coherence. In F. Huber and C. Schmidt-Petri (eds.), *Degrees of Belief*. Springer.

Joyce, J. 2011. The Development of Subjective Bayesianism. In D. Gabbay, S. Hartmann, and J. Woods (eds.), *Handbook of the History of Logic: Vol. 10, Inductive Logic*. Elsevier.

Kamp, H. 1967. The Treatment of 'Now' as a 1-Place Sentential Operator. multilith, UCLA.

Kanamori, A. 2007. Gödel and Set Theory. *Bulletin of Symbolic Logic*, 13:2.

Kanamori, A. 2008. Cohen and Set Theory. *Bulletin of Symbolic Logic*, 14:3.

Kanamori, A. 2009. *The Higher Infinite*, 2nd ed. Springer-Verlag.

Kanamori, A. 2012. Large Cardinals with Forcing. In D. Gabbay, A. Kanamori, and J. Woods (eds.), *Handbook of the History of Logic: Vol 6., Sets and Extensions in the Twentieth Century*. Elsevier.

Kanamori, A. 2012. Set Theory from Cantor to Cohen. In D. Gabbay, A. Kanamori, and J. Woods (eds.), *Handbook of the History of Logic: Vol. 6, Sets and Extensions in the Twentieth Century*. Elsevier.

Kant, I. 1787/1998. *Critique of Pure Reason*, tr. and ed. P. Guyer and A. Wood. Cambridge University Press.

Kaplan, D. 1975. How to Russell a Frege-Church. *Journal of Philosophy*, 72.

Kaplan, D. 1979. On the Logic of Demonstratives. *Journal of Philosophical Logic*, 8:1.

Kaplan, D. 1995. A Problem in Possible-World Semantics. In W. Sinnott-Armstrong, D. Raffman, and N. Asher (eds.), *Modality, Morality, and Belief: Essays in Honor of Ruth Barcan Marcus*. Cambridge University Press.

Katz, J. 1998. *Realistic Rationalism*. MIT Press.

Keisler, H.J. 1985. Probability Quantifiers. In J. Barwise and S. Feferman (op. cit.).

Kennedy, J. 2013. On Formalism Freeness: Implementing Gödel's 1946 Princeton Bicentennial Lecture. *Bulletin of Symbolic Logic*, 19:3.

Kleene, S. 1952. *Introduction to Metamathematics*. Van Nostrand.

Kment, B. 2012. Haecceitism, Chance, and Counterfactuals. *Philosophical Review*, 121:4.

Knuuttila, S. 1981. The Emergence of Deontic Logic in the Fourteenth Century. In Hilpinen (1981).

Koellner, P. 2009. On Reflection Principles. *Annals of Pure and Applied Logic*, 157.

Koellner, P. 2010. On Strong Logics of First and Second Order. *Bulletin of Symbolic Logic*, 16:1.

Koellner, P. 2013. Large Cardinals and Determinacy. *Stanford Encyclopedia of Philosophy*.

Koellner, P., and W.H. Woodin. 2010. Large Cardinals from Determinacy. In M. Foreman and A. Kanamori (eds.), *Handbook of Set Theory, Volume 3*. Springer.

Kolodny, N., and J. MacFarlane. 2010. Ifs and Oughts. *Journal of Philosophy*, 57:3.

Koopman, B. 1940. The Axioms and Algebra of Intuitive Probability. *Annals of Mathematics*, Second Series, 41:2.

Kratzer, A. 1977/2012. What 'Must' and 'Can' Must and Can Mean. *Linguistics and Philosophy*, 1:3. Reprinted in Kratzer, *Modals and Conditionals*. Oxford University Press.

Kratzer, A. 1979. Conditional Necessity and Possibility. In R. Bäuerle, U. Egli, and A.v. Stechow (eds.), *Semantics from Different Points of View*. Springer.

Kratzer, A. 1981/2012. Partition and Revision: the Semantics of Counterfactuals. *Journal of Philosophical Logic*, 10:2. Reprinted in Kratzer (2012).

Kremer, P. 2009. Dynamic Topological S5. *Annals of Pure and Applied Logic*, 160.

Kriegel, U. 2015. *The Varieties of Consciousness*. Oxford University Press.

Kripke, S. 1963. Semantical Considerations on Modal Logic. *Acta Philosophica Fennica*, 16.

Kripke, S. 1965. Semantical Analysis of Intuitionistic Logic I. In J. Crossley and M. Dummett (eds.), *Formal Systems and Recursive Functions*. North Holland.

Kripke, S. 1980. *Naming and Necessity*. Harvard University Press.

Kripke, S. 2005. Russell's Notion of Scope. *Mind*, 114.

Kuhn, S. 1980. Quantifiers as Modal Operators. *Studia Logica*, 39:2-3.

Kurz, A., and A. Palmigiano. 2013. Epistemic Updates on Algebras. *Logical Methods in Computer Science*, 9:4:17.

Lando, T. 2010. Completeness of S4 with Respect to the Lebesque Measure Algebra based on the Unit Interval. *Journal of Philosophical Logic*, DOI 10.1007/s10992-010-9161-3.

Lando, T. 2015. First Order S4 and Its Measure-theoretic Semantics. *Annals of Pure and Applied Logic*, 166.

Laplace, P. 'Memoir on the Probability of the Causes of Events', tr. S. Stigler. *Statistical Science*, 1:3.

Lawvere, F.W. 2005. An Elementary Theory of the Category of Sets. *Reprints in Theory and Applications of Categories*, 11.

Leibniz, G.W. 1704/1981. *New Essays on Human Understanding*, tr. and ed. P. Remnant and J. Bennett. Cambridge University Press.

Leibniz, G.W. 1714/1998. *Monadology*, tr. R. Franks and R.S. Woolhouse. In Franks and Woolhouse (eds.), *G.W. Leibniz: Philosophical Texts*. Oxford University Press.

Leitgeb, H. 2009. On Formal and Informal Provability. In O. Bueno and Ø. Linnebo (eds.), *New Waves in Philosophy of Mathematics*. Palgrave Macmillan.

Leitgeb, H., and R. Pettigrew. 2010. An Objective Justification of Bayesianism I: Measuring Inaccuracy. *Philosophy of Science*, 77.

Lemmon, E.J. (with D. Scott). 1966/1977. Intensional Logic/*An Introduction to Modal Logic*, ed. K. Segerberg. In N. Rescher (ed.), *American Philosophical Quarterly Monograph Series*, Vol. 11. Basil Blackwell.

Leng, M. 2007. What's There to Know? In Leng, A. Paseau, and M. Potter (eds.), *Mathematical Knowledge*. Oxford University Press.

Leng, M. 2009. 'Algebraic' Approaches to Mathematics. In O. Bueno and Ø. Linnebo (eds.), *New Waves in Philosophy of Mathematics*. Palgrave Macmillan.

Leng, M. 2010. *Mathematics and Reality*. Oxford University Press.

Levine, J. 1983. Materialism and Qualia: The Explanatory Gap. *Pacific Philosophical Quarterly*, 64.

Lewis, C.I. 1923. A Pragmatic Conception of the Apriori. *Journal of Philosophy*, 20:7.

Lewis, D. 1969. Review of *Art, Mind, and Religion. Journal of Philosophy*, 66:1.

Lewis, D. 1972. Psychophysical and Theoretical Identifications. *Australasian Journal of Philosophy*, 50:3.

Lewis, D. 1973. *Counterfactuals*. Blackwell.

Lewis, D. 1979. Attitudes De Dicto and De Se. *Philosophical Review*, 88:4.

Lewis, D. 1980. Index, Context, and Content. In S. Kanger and S. Öhman (eds.), *Philosophy and Grammar*. D. Reidel Publishing. Reprinted in Lewis (1998), *Papers in Philosophical Logic*. Cambridge University Press.

Lewis, D. 1980/1987. A Subjectivist's Guide to Objective Chance. In Lewis (1987), *Philosophical Papers, Volume 2*. Oxford University Press.

Lewis, D. 1986. *On the Plurality of Worlds*. Basil Blackwell.

Lewis, D. 1981/1998. Ordering Semantics and Premise Semantics for Counterfactuals. {Journal of Philosophical Logic, 10. Reprinted in Lewis, *Papers in Philosophical Logic*. Cambridge University Press.

Lewis, D. 1988/1998. Statements Partly about Observation. *Philosophical Papers*, 17:1. Reprinted In Lewis (1998).

Lewis, D. 1999. New Work for a Theory of Universals. In Lewis (op. cit.), *Papers in Metaphysics and Epistemology*. Cambridge University Press.

Lewis, D. 1999. *Papers in Metaphysics and Epistemology*. Cambridge University Press.

Lim, S., S. Padmala, and L. Pessoa. 2008. Affective Learning Modulates Spatial Competition During Low-load Attentional Conditions. *Neuropsychologica*, 46.

Linnebo, Ø. 2006. Sets, Properties, and Unrestricted Quantification. In A. Rayo and G. Uzquiano (eds.), *Absolute Generality*. Oxford University Press.

Linnebo, Ø. 2007. Burgess on Plural Logic and Set Theory. *Philosophia Mathematica*, 3:15.

Linnebo, Ø. 2008. Structuralism and the Notion of Dependence. *Philosophical Quarterly*, 58:230.

Linnebo, Ø. 2009. Bad Company Tamed. *Synthese*, 170.

Linnebo, Ø. 2010. Some Criteria on Acceptable Abstraction. *Notre Dame Journal of Formal Logic*, 52:3.

Linnebo, Ø. 2010. Pluralities and Sets. *Journal of Philosophy*, 107:3.

Linnebo, Ø. 2013. The Potential Hierarchy of Sets. *Review of Symbolic Logic*, 6:2.

Linnebo, Ø., and R. Pettigrew. 2014. Two Types of Abstraction for Structuralism. *Philosophical Quarterly*, doi: 10.1093/pq/pqt044.

Linnebo, Ø, and A. Rayo. 2012. Hierarchies, Ontological and Ideological. *Mind*, 121:482.

Linnebo, Ø., and G. Uzquiano. 2009. What Abstraction Principles Are Acceptable? Some Limitative Results. *British Journal for the Philosophy of Science*, 60.

Liu, T., J. Abrams, and M. Carrasco. 2009. Voluntary Attention Enhances Contrast Appearance. *Psychological Science*, 20:3.

Löb, M.H. 1955. Solution of a Problem of Leon Henkin. *Journal of Symbolic Logic*, 20:2.

MacFarlane, J. 2002. Frege, Kant, and the Logic of Logicism. *Philosophical Review*, 111:1.

MacFarlane, J. 2005. Making Sense of Relative Truth. *Proceedings of the Aristotelian Society*, 105.

Machery, E. 2005. You Don't Know How You Think: Introspection and the Language of Thought. *British Journal for the Philosophy of Science*, 56:3.

Machery, E. 2006. Two Dogmas of Neo-Empiricism. *Philosophy Compass*, 1/4.

Machery, E. 2009. *Doing without Concepts.* Oxford University Press.

Mackie, P. 2006. *How Things Might Have Been: Individuals, Kinds, and Essential Properties.* Oxford University Press.

Macpherson, F. 2011. Individuating the Senses. In Macpherson (ed.), *The Senses.* Oxford University Press.

Maddy, P. 1988,a. Believing the Axioms I. *Journal of Symbolic Logic*, 53:2.

Maddy, P. 1988,b. Believing the Axioms II. *Journal of Symbolic Logic*, 53:3.

Maddy, P. 1990. *Realism in Mathematics.* Oxford University Press.

Malament, D. ms. Notes on Geometry and Spacetime.

Mamassian, P., M. Landy, and L. Maloney. 2002. Bayesian Modelling of Visual Perception. In R. Rao and M. Lewicki (eds.), *Probabilistic Models of the Brain.* MIT Press.

Marcus, G. 2001. *The Algebraic Mind: Integrating Connectionism and Cognitive Science.* MIT Press.

Maudlin, T. 2010. Time, Topology, and Physical Geometry. *Proceedings of the Aristotelian Society Supplementary Volume,* 84.

Maudlin, T. 2013. The Nature of the Quantum State. In A. Ney and D. Albert (eds.), *The Wave Function.* Oxford University Press.

McCarthy, T. 1981. The Idea of a Logical Constant. *Journal of Philosophy,* 78:9.

McCarthy, T. 1987. Modality, Invariance, and Logical Truth. *Journal of Philosophical Logic,* 16:4.

McDowell, J. 1994. *Mind and World.* Harvard University Press.

McFetridge, I. 1990. *Logical Necessity,* eds. J. Haldane and R. Scruton. Aristotelian Society Series, Volume 11.

McGee, V. 1991. *Truth, Vagueness, and Paradox.* Hackett Publishing.

McGee, V. 1992. Two Problems with Tarski's Theory of Logical Consequence. *Proceedings of the Aristotelian Society,* New Series, 92.

McGee, V. 1996. Logical Operations. *Journal of Philosophical Logic,* 25:6.

McKinsey, J., and A. Tarski. 1944. The Algebra of Topology. *The Annals of Mathematics, Second Series,* 45:1.

McLarty, C. 2008. Introduction to Categorical Foundations of Mathematics.

McLarty, C. 2009. What Does It Take to Prove Fermat's Last Theorem? Gorthendieck and the Logic of Number Theory. *Bulletin of Symbolic Logic.*

Menger, K. 1930/1979. On Intuitionism. In Menger, *Selected Papers in Logic and Foundations, Didactics, Economics,* ed. H. Mulder. D. Reidel Publishing.

Meyer, J.J., and W. van der Hoek. *Epistemic Logic for AI and Computer Science.* Cambridge University Press.

Milne, P. ms. Measuring Confirmation.

Mole, C., D. Smithies, and W. Wu (eds.). 2011. *Attention: Philosophical and Psychological Essays*. Oxford University Press.

Montague, R. 1960/1974. Logical Necessity, Physical Necessity, Ethics, and Quantifiers. In Montague, *Formal Philosophy*, ed. R. Thomason. Yale University Press.

Montague, R. 1963. Syntactical Treatments of Modality, with Corollaries on Reflexion Principles and Finite Axiomatizability. *Acta Philosophica Fennica*, 16.

Montero, B. 2010. A Russellian Response to the Structural Argument against Physicalism. *Journal of Consciousness Studies*, 17:3-4.

Moore, G.E. c.1941-1942/1962. *Commonplace Book, 1919-1953*, ed. C. Lewy. George Allen and Unwin.

Moschovakis, Y. 1998. On the Founding of a Theory of Algorithms. In H. Dales and G. Olivieri (eds.), *Truth in Mathematics*. Clarendon Press.

Moss, S. 2011. Scoring Rules and Epistemic Compromise. *Mind*, 120:480.

Moss, S. 2015. Time-slice Epistemology and Action under Indeterminacy. In J. Hawthorne and T. Gendler, *Oxford Studies in Epistemology, Volume 5*. Oxford University Press.

Moss, S. 2015. On the Semantics and Pragmatics of Epistemic Vocabulary. *Semantics and Pragmatics*, 8(5), 1?81.

Moss, S. 2018. *Probabilistic Knowledge*. Oxford University Press.

Myhill, J. 1985. Intensional Set Theory. In Shapiro (ed.), *Intensional Mathematics*. Elsevier Science Publishers.

Nagel, J. 2007. Epistemic Intuitions. *Philosophy Compass*, 2/6.

Nagel, J. 2013a. Defending the Evidential Value of Epistemic Intuitions. *Philosophy and Phenomenological Research*, 87:1.

Nagel, J. 2013b. Knowledge as a Mental State. In T. Gendler and J. Hawthorne (eds.), *Oxford Studies in Epistemology, Volume 4*. Oxford University Press.

Nagel, J. 2014. Intuition, Reflection, and the Command of Knowledge. *Proceedings of the Aristotelian Society*, 88.

Neitz, J., and G. Jacobs. 1986. Polymorphism of the Long-wavelength Cone in Normal Human Colour Vision. *Nature*, 323.

Neitz, M., and J. Neitz. 2000. Molecular Genetics of Color Vision and Color Vision Defects. *Archives of Ophthalmology*, 118.

Newell, A. 1980. Physical Symbol Systems. *Cognitive Science*, 4.

Ney, A. 2013. Ontological Reduction and the Wave Function Ontology. In Ney and D. Albert (eds.), *The Wave Function.* Oxford University Press.

Nichols, S. (ed.). 2006. *The Architecture of the Imagination.* Oxford University Press.

Nida-Rümelin, M. 2007. Grasping Phenomenal Properties. In T. Alter and S. Walter (eds.), *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism.* Oxford University Press.

Nolan, D. 1996. Recombination Unbound. *Philosophical Studies*, 84:2-3.

Nolan, D. 2002. *Topics in the Philosophy of Possible Worlds.* Routledge.

North, J. 2013. The Structure of a Quantum World. In D. Albert and A. Ney (eds.), *The Wavefunction.* Oxford University Press.

Nozick, R. 1981. *Philosophical Explanations.* Harvard University Press.

O'Keefe, J., and L. Nadel. 1978. *The Hippocampus as a Cognitive Map.* Oxford University Press.

Oliver, A., and T. Smiley. 2013. *Plural Logic.* Oxford University Press.

Pacherie, E. 2012. Action. In K. Frankish and W. Ramsey (eds.), *The Cambridge Handbook of Cognitive Science.* Cambridge University Press.

Padmala, S., and L. Pessoa. 2008. Affective Learning Enhances Visual Detection and Responses in Primary Visual Cortex. *Journal of Neuroscience*, 9.

Papineau, D. 2015. Can We Really See a Million Colors? In P. Coates and S. Coleman (eds.), *Phenomenal Qualities.* Oxford University Press.

Paris, J.B. 2001. A Note on the Dutch Book Method. In *Proceedings of the Second International Symposium on Imprecise Probabilities and Their Applications, ISIPTA.* Shaker.

Parsons, C. 1964. Frege's Theory of Number. In M. Black (ed.), *Philosophy in America.* Allen and Unwin.

Parsons, C. 1977/1983. What is the Iterative Conception of Set? In P. Benacerraf and H. Putnam (eds.), *Philosophy of Mathematics*, 2nd ed. Cambridge University Press.

Parsons, C. 1979-1980. Mathematical Intuition. *Proceedings of the Aristotelian Society*, New Series, 80.

Parsons, C. 1983. *Mathematics in Philosophy*. Cornell University Press.

Parsons, C. 1993. On Some Difficulties concerning Intuition and Intuitive Knowledge. *Mind*, 102:406.

Parsons, C. 1997. What Can We Do 'In Principle'? In M. Chiara, K. Doets, D. Mundici, and J. van Benthem (eds.), *Logic and Scientific Methods*. Kluwer Academic Publishers.

Parsons, C. 2008. *Mathematical Thought and Its Objects*. Cambridge University Press.

Pascal, B. 1654/1959. 'Fermat and Pascal on Probability', tr. V. Sanford. In D. Smith (ed.), *A Source Book in Mathematics*. Dover.

Peacocke, C. 1992. *A Study of Concepts*. MIT Press.

Peacocke, C. 2003. *The Realm of Reason*. Oxford University Press.

Peacocke, C. 2008. *Truly Understood*. Oxford University Press.

Peano, G. 1889/1967. The Principles of Arithmetic, Presented by a New Method (tr. J. van Heijenoort). In van Heijenoort (1967).

Pearl, J. 2000. *Causality*. Cambridge University Press.

Peirce, C.S. 1933. *Collected Papers of Charles Sanders Peirce: Vol. IV, The Simplest Mathematics*, ed., C. Hartshorne and P. Weiss. Harvard University Press.

Perry, J. 2001. *Knowledge, Possibility, and Consciousness*. MIT Press.

Pettigrew, R. 2012. Accuracy, Chance, and the Principal Principle. *Philosophical Review*, 121:2.

Pettigrew, R. 2014. Accuracy and Evidence. *Dialectica*, 67:4.

Pettigrew, R. ms. Categorical Semantics for First-order Logic.

Pettigrew, R. ms. An Introduction to Toposes.

Piccinini, G. 2004. Functionalism, Computationalism, and Mental States. *Studies in History and Philosophy of Science*, 35.

Pitt, D. 2004. The Phenomenology of Cognition. *Philosophy and Phenomenological Research*, 69:1.

Pollard, C. 2013. Agnostic Hyperintensional Semantics. *Synthese*, DOI 10.1007/s11229-013-0373-2.

Potter, M. 2009. The Logic of the *Tractatus*. In D. Gabbay and J. Woods (eds.), *Handbook of the History of Logic: Vol. 5, Logic from Russell to Church.* Elsevier.

Pratt, V. 1976. Semantical Considerations on Floyd-Hoare Logic. In *Proceedings of the 17th Annual Symposium on Foundations of Computer Science.* IEEE.

Price, H. 2013. *Expressivism, Pragmatism and Representationalism.* Cambridge University Press.

Priest, G. 2006. *In Contradiction*, 2nd ed. Oxford University Press.

Priest, G. 2009. Beyond the Limits of Knowledge. In Salerno (2009).

Priest, G. 2014. Revising Logic. In P. Rush (ed.), *The Metaphysics of Logic.* Cambridge University Press.

Priest, G. ms/slides. How Do You Know What the Right Logic Is? http://fitelson.org/bridges2/priest.pdf. Downloaded 10/18/2015.

Prinz, J. 2002. *Furnishing the Mind.* MIT Press.

Prinz, J. 2012. *The Conscious Brain.* Oxford University Press.

Pryor, J. 1999. Immunity to Error through Misidentification. *Philosophical Topics*, 26.

Putnam, H. 1967. Mathematics without Foundations. *Journal of Philosophy*, 64.

Putnam, H. 1967. Psychological Predicates. In W. Capitan and D. Merrill (eds.) *Art, Mind, and Religion.* University of Pittsburgh Press. Reprinted as 'The Nature of Mental States', in N. Block (ed.), *Readings in Philosophy of Psychology, Volume One.* Harvard University Press, 1990.

Putnam, H. 1980. Models and Reality. *Journal of Symbolic Logic*, 45:3.

Putnam, H. 1981. *Reason, Truth, and History.* Cambridge University Press.

Pylyshyn, Z. 1978. Computational Models and Empirical Constraints. *Brain and Behavioral Sciences*, 1.

Pylyshyn, Z. 2007. *Things and Places.* MIT Press.

Quine, W.V. 1951. Two Dogmas of Empiricism. *Philosophical Review*, 60:1.

Quine, W.V. 1963. Necessary Truth. In Quine (1976), *The Ways of Paradox*, 2nd ed. Harvard University Press.

Quine, W.V. 1968. Propositional Objects. *Crtica*, 2:5.

Raffman, D. 2000. Is Perceptual Indiscriminability Non-transitive? In C. Hill (ed.), *Philosophical Topics*, 28.

Ramsey, F.P. 1926. Truth and Probability. In Ramsey (1960), *The Foundations of Mathematics*, ed. R. Braithwaite. Littlefield, Adams, and Co.

Rao, A., and M. Georgeff. 1991. Modeling Rational Agents within a BDI-Architecture. In J. Allen, R. Fikes, and E. Sandewall (eds.), *KR '91: Principles of Knowledge Representation and Reasoning*. Morgan Kaufmann.

Rasiowa, H. 1963. On Modal Theories. *Acta Philosophica Fennica*, 16.

Rayo, A. 2006. Beyond Plurals. In Rayo and G. Uzquiano (eds.), *Absolute Generality*. Oxford University Press.

Rayo, A. 2013. *The Construction of Logical Space*. Oxford University Press.

Rayo, A., and T. Williamson. 2003. A Completeness Theorem for Unrestricted First-Order Languages. In JC Beall (ed.), *Liars and Heaps*. Oxford University Press.

Read, S. 2009. Plural Signification and the Liar Paradox. *Philosophical Studies*, 145.

Read, S. 2010. Field's Paradox and Its Medieval Solution. *History and Philosophy of Logic*, 31.

Recanati, F. 2007. *Perspectival Thought*. Oxford University Press.

Reinhardt, W. 1974. Remarks on Reflection Principles, Large Cardinals, and Elementary Embeddings. In T. Jech (ed.), *Proceedings of Symposia in Pure Mathematics, Vol. 13, Part 2: Axiomatic Set Theory*. American Mathematical Society.

Reinhardt, W. 1986. Epistemic Theories and the Interpretation of Gödel's Incompleteness Theorems. *Journal of Philosophical Logic*, 15:4.

Reinhardt, W. 1988. Epistemic Set Theory. *Notre Dame Journal of Formal Logic*, 29:2.

Rescorla, M. 2009. Cognitive Maps and the Language of Thought. *British Journal for the Philosophy of Science*, 60.

Rescorla, M. 2013. Bayesian Perceptual Psychology. In M. Mohan (ed.), *The Oxford Handbook of the Philosophy of Perception*. Oxford University Press.

Rescorla, M. 2015. The Representational Foundations of Computation. *Philosophia Mathematica*, 23:3.

Restall, G. 2009. Not Every Truth Can Be Known. In Salerno (2009).

Reynolds, J., and D. Heeger. 2009. The Normalization Model of Attention. *Neuron* 61.2.

Rittberg, C. 2015. How Woodin Changed His Mind: New Thoughts on the Continuum Hypothesis. *Archive for History of Exact Sciences*, 69:2.

Roca-Royes, S. 2011. Conceivability and *De Re* Modal Knowledge. *Nous*, 45:1.

Roca-Royes, S. 2011. Modal Knowledge and Counterfactual Knowledge. *Logique Et Analyse*, 54:216.

Roca-Royes, S. 2012. Essentialist Blindness Would Not Preclude Counterfactual Knowledge. *Philosophia Scientiae*, 16:2.

Roca-Royes, S. 2016. Similarity and Possibility. In B. Fischer and F. Leon (eds.), *Modal Epistemology after Rationalism*. Synthese Library.

Rosen, G. 2010. Metaphysical Dependence: Grounding and Reduction. In B. Hale and A. Hoffman (eds.), *Modality: Metaphysics, Logic, and Epistemology*. Oxford University Press.

Rosen, G., and S. Yablo. 2020. Solving the Caesar Problem – With Metaphysics. In A. Miller (ed.), *Logic, Language, and Mathematics: Themes from the Philosophy of Crispin Wright*. Oxford University Press.

Routley, R. 1981/2010. Necessary Limits to Knowledge. In M. Edgar, N. Otto, and Z. Gerhard (eds.), *Essays in Scientific Philosophy*. Comes Verlag. Reprinted in *Synthese*, 173.

Routley, R., and R. Meyer. 1972,a. The Semantics of Entailment II. *Journal of Philosophical Logic*, 1.

Routley, R., and R. Meyer. 1972,b. The Semantics of Entailment III. *Journal of Philosophical Logic*, 1.

Routley, R., and R. Meyer. 1973. The Semantics of Entailment I. In H. Leblanc (ed.), *Truth, Syntax, and Modality*. North-Holland.

Ruetsche, L. 2011. *Interpreting Quantum Theories*. Oxford University Press.

Rumfitt, I. 2014. Truth and Meaning. *Proceedings of the Aristotelian Society Supplementary Volume*, 88.

Russell, B. 1903/2010. *The Principles of Mathematics.* Routledge.

Russell, B. 1919. The Philosophy of Logical Atomism. *The Monist*, 29:3.

Russell, G. 2015. The Justification of the Basic Laws of Logic. *Journal of Philosophical Logic*, DOI 10.1007/s10992-015-9360-z.

Russell, J. 2013. Possible Worlds and the Objective World. *Philosophy and Phenomenological Research*, doi: 10.1111/PHPR.12052.

Russell, J. ms. Quality and Quantifiers.

Sagi, G. 2015. The Modal and Epistemic Arguments against the Invariance Criterion for Logical Terms. *Journal of Philosophy*, 112:3.

Salmon, N. 2010. Lambda in Sentences with Designations: An Ode to Complex Predication. *Journal of Philosophy*, 107:9.

Savage, L. 1954/1972. *The Foundations of Statistics.* Dover.

Segerberg, K. 1973. Two-dimensional Modal Logic. *Journal of Philosophical Logic*, 2:1.

Scedrov, A. 1985. Extending Gödel's Modal Interpretation to Type Theory and Set Theory. In Shapiro (ed.), *Intensional Mathematics.* Elsevier Science Publishers.

Schaffer, J. 2009. On What Grounds What. In D. Manley, D. Chalmers, and R. Wasserman (eds.), *Metametaphysics.* Oxford University Press.

Schaffer, J. 2009. Spacetime, the One Substance. *Philosophical Studies*, 145.

Schaffer, J. 2012. Grounding, Transitivity, and Contrastivity. In F. Correia and B. Schnieder (eds.), *Metaphysical Grounding.* Cambridge University Press.

Scharp, K. 2013. *Replacing Truth.* Oxford University Press.

Schechter, J. 2010. The Reliability Challenge and the Epistemology of Logic. *Philosophical Perspectives*, 24.

Scholl, B., and J.I. Flombaum. 2010. Object Persistence. In B. Goldstein (ed.), *Encyclopedia of Perception, Vol. 2.* SAGE Publications.

Schroeder, M. 2011. Ought, Agents, and Actions. *Philosophical Review*, 120:1.

Schwitzgebel, E. 2011. *Perplexities of Consciousness.* Bradford.

Seager, W. 1995. Consciousness, Information, and Panpsychism. *Journal of Consciousness Studies*, 2:3.

Segerberg, K. 1971. *An Essay in Classical Modal Logic. Filosofiska Studier*, Vol. 13. Uppsala Universitet.

Setiya, K. 2007. *Reasons without Rationalism.* Princeton University Press.

Setiya, K. 2014. Intention. *Stanford Encyclopedia of Philosophy.*

Shapiro, S. 1985. Epistemic and Intuitionistic Arithmetic. In Shapiro (ed.), *Intensional Mathematics.* Elsevier Science Publishers.

Shapiro, S. 1991. *Foundations without Foundationalism.* Oxford University Press.

Shapiro, S. 1994. Reasoning, Logic, and Computation. *Philosophia Mathematica*, 3:2.

Shapiro, S. 1997. *Philosophy of Mathematics.* Oxford University Press.

Shapiro, S. 1998. Logical Consequence: Models and Modality. In M. Schirn (ed.), *The Philosophy of Mathematics Today.* Oxford University Press.

Shapiro, S. 2000. Frege meets Dedekind: A Neologicist Treatment of Real Analysis. *Notre Dame Journal of Formal Logic*, 41:4.

Shapiro, S., and G. Hellman. 2015. Frege meets Aristotle: Points as Abstracts. *Philosophia Mathematica*, doi:10.1093/philmat/nkv021.

Shapiro, S., and Ø. Linnebo. 2015. Frege meets Brouwer (or Heyting or Dummett). *Review of Symbolic Logic*, doi:10.1017/S1755020315000039.

Shapiro, S., and A. Weir. 1999. New V, ZF, and Abstraction. *Philosophia Mathematica*, 3:7.

Shapiro, S., and C. Wright. 2006. All Things Indefinitely Extensible. In A. Rayo and G. Uzquiano (eds.), *Absolute Generality.* Oxford University Press.

Shea, N. 2013. Neural Mechanisms of Decision-making and the Personal Level. In K.W.M. Fulford, M. Davies, R. Gipps, G. Graham, J. Sadler, G. Stanghellini, and T. Thornton (eds.), *The Oxford Handbook of Philosophy and Psychiatry.* Oxford University Press.

Sher, G. 1991. *The Bounds of Logic.* MIT Press.

Sher, G. 2001. The Formal-Structural View of Logical Consequence. *Philosophical Review*, 110:2.

Sher, G. 2008. Tarski's Thesis. In D. Patterson (ed.), *New Essays on Tarski and Philosophy*. Oxford University Press.

Shoemaker, S. 1968. Self-reference and Self-awareness. *Journal of Philosophy*, 65.

Sider, T. 2007. Neo-Fregeanism and Quantifier Variance. *Proceedings of the Aristotelian Society Supplementary Volume*, 81.

Sider, T. 2009. Williamson's Many Necessary Existents. *Analysis*, 69.

Sider, T. 2011. *Writing the Book of the World*. Oxford University Press.

Siegel, S. 2010. *The Contents of Visual Experience*. Oxford University Press.

Sider, T. 2016. On Williamson and Simplicity in Modal Logic. *Canadian Journal of Philosophy*, 46:4-5.

Simons, P. 2016. Applications of Complex Numbers and Quaternions. In P. Ebert and M. Rossberg (eds.), *Abstractionism: Essays in Philosophy of Mathematics*. Oxford University Press.

Simpson, S. 1999. *Subsystems of Second Order Arithmetic*. Springer.

Skiles, A. 2015. Against Grounding Necessitarianism. *Erkenntnis*, 80.

Skow, B. 2005. *Once upon a Spacetime*. Ph.D. Dissertation, New York University.

Skow, B. 2008. Haecceitism, Anti-Haecceitism and Possible Worlds. *Philosophical Quarterly*, 58.

Smart, J.J.C. 1959. Sensations and Brain Processes. *Philosophical Review*, 68:2.

Smiley, T. 1963. The Logical Basis of Ethics. *Acta Philosophica Fennica*, 16.

Smithies, D. 2013. The Nature of Cognitive Phenomenology. *Philosophy Compass*, 8:8.

Smithies, D. 2013. The Significance of Cognitive Phenomenology. *Philosophy Compass*, 8:8.

Sorensen, R. 1988. *Blindspots*. Oxford University Press.

Sorensen, R. 1999. *Thought Experiments*. Oxford University Press.

Spelke, E., and K. Kinzler. 2007. Core Knowledge. *Developmental Science*, 10:1.

Stalnaker, R. 1968. A Theory of Conditionals. In J. Cornman (ed.), *Studies in Logical Theory*. Blackwell Publishing.

Stalnaker, R. 1978. Assertion. In P. Cole (ed.), *Syntax and Semantics, Vol. 9*. Academic Press.

Stalnaker, R. 1983. *Inquiry*. MIT Press.

Stalnaker, R. 2003. *Ways a World Might Be*. Oxford University Press.

Stalnaker, R. 2004. Assertion Revisited. *Philosophical Studies*, 118.

Stalnaker, R. 2008. *Our Knowledge of the Internal World*. Oxford University Press.

Stalnaker, R. 2011. Conditional Propositions and Conditional Assertions. In A. Egan and B. Weatherson (eds.), *Epistemic Modality*. Oxford University Press.

Stalnaker, R. 2012. *Mere Possibilities*. Princeton University Press.

Stanley, J., and T. Williamson. 2001. Knowing How. *Journal of Philosophy*, 98.

Starr, W. 2012. Epistemic Modality, Contextualism, and Relativism.

Starr, W. 2013. A Uniform Theory of Conditionals. *Journal of Philosophical Logic*, http://dx.doi.org/10.1007/s10992-013-9300-8.

Starr, W. 2016. Dynamic Expressivism about Deontic Modality. In N. Charlow and M. Chrisman (eds.), *Deontic Modality*. Oxford University Press.

Stoljar, D. 2001. Two Conceptions of the Physical. *Philosophy and Phenomenological Research*, 62:2.

Stoljar, D. 2014. Four Kinds of Russellian Monism. In U. Kriegel (ed.), *Contemporary Debates in Philosophy of Mind*. Routledge.

Strawson, G. 2006. Realistic Monism. *Journal of Consciousness Studies*, 13.

Studd, J. 2013. The Iterative Conception of Set: a (Bi-)Modal Characterisation. *Journal of Philosophical Logic*, 42.

Takeuchi, M. 1985. Topological Coalgebras. *Journal of Algebra*, 97.

Tarski, A. 1930/1983. On Some Fundamental Concepts of Metamathematics. In Tarski (1983), *Logic, Semantics, and Metamathematics* (2nd ed.), tr. J.H. Woodger. Hackett Publishing.

Tarski, A. 1936/1983. On the Concept of Logical Consequence. In Tarski (op. cit.).

Thomason, R. 1981. Deontic Logic and the Role of Freedom in Moral Deliberation. In Hilpinen (1981).

Thomasson, A. 2007. Modal Normativism and the Methods of Metaphysics. *Philosophical Topics*, 35:1-2.

Timpson, C. 2013. *Quantum Information Theory and the Foundations of Quantum Mechanics.* Oxford University Press.

Treisman, A., and G. Gelade. 1980. A Feature-integration Theory of Attention. *Cognitive Psychology*, 12:16.

Tsuchiya, N., N. Block and C. Koch. 2012. Top-down Attention and Consciousness: comment on Cohen, et.al. *Trends in Cognitive Sciences*, 16:11.

Turing, A. 1936. On Computable Numbers, with an Application to the *Entscheidungsproblem*. *Proceedings of the London Mathematical Society*, 2:42

Turing, A. 1950. Computing Machinery and Intelligence. *Mind*, 59: 236.

Turner, J. 2016. *The Facts in Logical Space.* Oxford University Press.

Tye, M. 2000. *Consciousness, Color, and Content.* MIT Press.

The Univalent Foundations Program. *Homotopy Type Theory: Univalent Foundations of Mathematics.* Institute for Advanced Study.

Uzquiano, G. 2003. Plural Quantification and Classes. *Philosophia Mathematica*, 3:11.

Uzquiano, G. 2015,a. Varieties of Indefinite Extensibility. *Notre Dame Journal of Formal Logic*, 58:1.

Uzquiano, G. 2015,b. Modality and Paradox. *Philosophy Compass*, 10:4.

Uzquiano, G. 2015,c. Recombination and Paradox. *Philosophers' Imprint*, 15:19.

van Fraassen, B. 1969. Facts and Tautological Entailments. *Journal of Philosophy*, 66.

van Heijenoort, J. (ed.). 1967. *From Frege to Gödel.* Harvard University Press.

Väyrynen, P. 2013. Grounding and Normative Explanation. *Aristotelian Society Supplementary Volume*, 87:1.

Velleman, J.D. 1989. *Practical Reflection*. Princeton University Press.

Veltman, F. 1996. Defaults in Update Semantics. *Journal of Philosophical Logic*, 25:3.

Venema, Y. 2007. Algebras and Coalgebras. In P. Blackburn, J. van Benthem, and F. Wolter (eds.), *Handbook of Modal Logic*. Elsevier.

Venema, Y. 2012. Lectures on the Modal μ-Calculus.

Venema, Y. 2013. Cylindric Modal Logic. In H. Andrka, M. Ferenczi, and I. Nmeti (eds.), *Cylindric-Like Algebras and Algebraic Logic*. Jnos Bolyai Mathematical Society and Springer-Verlag.

Venema, Y. 2014. Expressiveness modulo Bisimilarity. In A. Baltag and S. Smets (eds.), *Johan van Benthem on Logic and Information Dynamics*. Springer.

Vetter, B. 2013. 'Can' without Possible Worlds: Semantics for Anti-Humeans. *Philosophers' Imprint*, 13:16.

Vlach, F. 1973. *'Now' and 'Then'*. Ph.D. Thesis, UCLA.

Voevodsky, V. 2006. A Very Short Note on the Homotopy λ-Calculus. Unpublished.

von Neumann, J. 1923/1967. On the Introduction of Transfinite Numbers (tr. van Heijenoort). In van Heijenoort (1967).

von Neumann. 1925/1967. An Axiomatization of Set Theory (tr. S. Bauer-Mengelberg and D. Follesdal). In J. van Heijenoort (ed.), *From Frege to Gödel*. Harvard University Press.

von Neumann, J., and O. Morgenstern. 1944. *Theory of Games and Economic Behavior*. Princeton University Press.

Wallace, D., and C. Timpson. 2010. Quantum Mechanics on Spacetime I: Spacetime State Realism. *British Journal for the Philosophy of Science*, 61.

Walsh, S. 2016. Fragments of Frege's *Grundgesetze* and Gödel's Constructible Universe. *Journal of Symbolic Logic*, 81:2.

Wang, H. 1974. *From Mathematics to Philosophy*. Routledge and Kegan Paul.

Wang, H. 1986. *A Logical Journey*. MIT Press.

Wansing, H. 2002. Diamonds are a Philosopher's Best Friends. *Journal of Philosophical Logic*, 31.

Waterlow, S. 1982. *Passage and Possibility: A Study of Aristotle's Modal Concepts.* Oxford University Press.

Watzl, S. Forthcoming. Can Intentionalism Explain how Attention Affects Appearances? In A. Pautz and D. Stoljar (eds.), *Themes from Block.* MIT Press.

Weir, A. 2003. Neo-Fregeanism: An Embarrassment of Riches. *Notre Dame Journal of Formal Logic*, 44.

Whittle, B. 2009. Epistemically Possible Worlds and Propositions. *Nous*, 43:2.

Williams, J.R.G. 2012. Gradational Accuracy and Nonclassical Semantics. *Review of Symbolic Logic*, 5:4.

Williams, J.R.G. 2014. Decision-making under Indeterminacy. *Philosophers' Imprint*, 14:4.

Williams, J.R.G. 2014. Nonclassical Minds and Indeterminate Survival. *Philosophical Review*, 123:4.

Williamson, T. 1990/2013. *Identity and Discrimination.* Wiley-Blackwell.

Williamson, T. 1994. *Vagueness.* Routledge.

Williamson, T. 1998. Indefinite Extensibility. *Grazer Philosophische Studien*, 55.

Williamson, T. 2002. *Knowledge and Its Limits.* Oxford University Press.

Williamson, T. 2007. *The Philosophy of Philosophy.* Oxford University Press.

Williamson, T. 2009. Some Computational Constraints in Epistemic Logic. In S. Rahman, J. Symons, D. Gabbay, and J. van Bendegem (eds.), *Logic, Epistemology, and the Unity of Science.* Springer.

Williamson, T. 2013. *Modal Logic as Metaphysics.* Oxford University Press.

Williamson, T. 2013. How Deep is the Distinction between A Priori and A Posteriori Knowledge? In A. Casullo and J. Thurow (eds.), *The Apriori in Philosophy.* Oxford University Press.

Williamson, T. 2014. Logic, Metalogic, and Neutrality. *Erkenntnis*, 79.

Williamson, T. 2014. A Note on Gettier Cases in Epistemic Logic. *Philosophical Studies*, DOI 10.1007/s11098-014-0357-1.

Williamson, T. 2014. Very Improbable Knowing. *Erkenntnis*, DOI 10.1007/s10670-013-9590-9.

Williamson, T. 2016. Knowing by Imagining. In A. Kind and P. Kung (eds.), *Knowledge through Imagination.* Oxford University Press.

Williamson, T. 2016. Absolute Provability and Safe Knowledge of Axioms. In L. Horsten and P. Welch (eds.), *Gödel's Disjunction.* Oxford University Press.

Williamson, T. 2016. Modal Science. *Canadian Journal of Philosophy*, 46:4-5.

Williamson, T. Forthcoming. Semantic Paradoxes and Abductive Methodology. In B. Armour-Garb (ed.), *The Relevance of the Liar.* Oxford University Press.

Williamson, T. Forthcoming. Dummett on the Relation between Logics and Metalogics. In M. Frauchiger and W.K. Essler (eds.), 'Volume on Michael Dummett in the Lauener Library of Analytical Philosophy'. De Gruyter.

Wittgenstein, L. 1921/1974. *Tractatus Logico-Philosophicus*, tr. D.F. Pears and B.F. McGuinness. Routledge.

Wittgenstein, L. 1933-1937/2005. *The Big Typescript: TS 213*, tr. and ed. C. Luckhardt and M. Aue. Blackwell Publishing.

Wittgenstein, L. 1969. *On Certainty*, tr. D. Paul and G.E.M. Anscombe, ed. Anscombe and G.H. von Wright. Basil Blackwell.

Wittgenstein, L. 1979. *Notebooks, 1914-1916*, 2nd ed., tr. G.E.M Anscombe, ed. G.H. von Wright and Anscombe. University of Chicago Press.

Wittgenstein, L. 2010. *Remarks on the Foundations of Mathematics*, 3rd ed., tr. G.E.M. Anscombe, ed. G.H. von Wright, R. Rees, and Anscombe. Basil Blackwell.

Woodin, W.H. 1999. *The Axiom of Determinacy, Forcing Axioms, and the Nonstationary Ideal.* de Gruyter.

Woodin, W.H. 2001. The Continuum Hypothesis, Part I. *Notices of the American Mathematical Society*; 48:6.

Woodin, W.H. 2001. The Continuum, Hypothesis, Part II. *Notices of the American Mathematical Society*; 48:7.

Woodin, W.H. 2010. Strong Axioms of Infinity and the Search for V. *Proceedings of the International Congress of Mathematicians.*

Woodin, W.H. 2011. The Realm of the Infinite. In M. Heller and Woodin (eds.), *Infinity: New Research Frontiers*. Cambridge University Press.

Woodin, W.H. ms. The Ω Conjecture.

Woodward, J., and C. Hitchcock. 2003. Explanatory Generalizations, Part I. *Noûs*, 37:1.

von Wright, G.H. 1957. On Double Quantification. In von Wright, *Logical Studies*. Routledge.

Wright, C. 1983. *Frege's Conception of Numbers as Objects*. Aberdeen University Press.

Wright, C. 1985. Skolem and the Sceptic. *Proceedings of the Aristotelian Society, Supplementary Volume*, 59.

Wright, C. 1987/1993. *Realism, Meaning, and Truth*. Blackwell.

Wright, C. 1995. Intuitionists Are Not (Turing) Machines. *Philosophia Mathematica*, 3:3.

Wright, C. 1997. The Philosophical Significance of Frege's Theorem. In R. Heck (ed.), *Language, Thought, and Logic*. Oxford University Press. Reprinted in Hale and Wright (2001).

Wright, C. 2000. Neo-Fregean Foundations for Real Analysis. *Notre Dame Jouranl in Formal Logic*, 41:4.

Wright, C. 2004. Warrant for Nothing (and Foundations for Free)? *Proceedings of the Aristotelian Society, Supplementary Volume*, 78:1.

Wright, C. 2004. Intuition, Entitlement and the Epistemology of Logical Laws. *Dialectica*, 58:1.

Wright, C. 2007. The Perils of Dogmatism. In S. Nuccetelli and G. Seay (eds.), *Themes from G.E. Moore*. Oxford University Press.

Wright, C. 2007. On Quantifying into Predicate Position. In M. Leng, A. Paseau, and M. Potter (eds.), *Mathematical Knowledge*. Oxford University Press.

Wright, C. 2012. Reflections on Recanati's 'Immunity to Error through Misidentification'. In S. Prosser and F. Recanati (eds.), *Immunity to Error through Misidentification*. Oxford University Press.

Wright, C. 2012. Replies, Part IV: Warrant Transmission and Entitlement. In A. Coliva (ed.), *Mind, Meaning and Knowledge*. Oxford University Press.

Wright, C. 2014. On Epistemic Entitlement II. In D. Dodd and E. Zardini (eds.), *Scepticism and Perceptual Justification*. Oxford University Press.

Wright, C. Frege and Logicism. 2020. In In A. Miller (ed.), *Logic, Language, and Mathematics: Themes from the Philosophy of Crispin Wright*. Oxford University Press.

Wu, W. 2011. Attention as Selection for Action. In C. Mole, D. Smithies, and W. Wu (eds.), *Attention: Philosophical and Psychological Essays*. Oxford University Press.

Yablo, S. 1993. Paradox without Self-reference. *Analysis*, 53.

Yablo, S. 2008. *Thoughts*. Oxford University Press.

Yablo, S. 2011. A Problem about Permission and Possibility. In A. Egan and B. Weatherson (eds.), *Epistemic Modality*. Oxford University Press.

Yablo, S. 2014. *Aboutness*. Oxford University Press.

Yalcin, S. 2007. Epistemic Modals. *Mind*, 116:464.

Yalcin, S. 2011. Nonfactualism about Epistemic Modality. In A. Egan and B. Weatherson (eds.), *Epistemic Modality*. Oxford University Press.

Yalcin, S. 2012. Bayesian Expressivism. *Proceedings of the Aristotelian Society*, doi: 10.1111/j.1467-9264.2012.00329.x

Yalcin, S. 2016. Beliefs as Question-sensitive. *Philosophy and Phenomenological Research*, doi: 10.1111/phpr.12330.

Zeimbekis, J., and A. Raftopoulos (eds.). 2015. *The Cognitive Penetrability of Perception*. Oxford University Press.

Zermelo, E. 1908/1967. A New Proof of the Possibility of a Well-ordering (tr. S. Bauer-Mengelberg). In van Heijenoort (1967).

Zermelo, E. 1930a/2010. On the Set-theoretic Model. In Zermelo (2010), *Collected Works, Volume 1*, tr. and ed. H.-D. Ebbinghaus, C. Fraser, and A. Kanamori. Springer.

Zermelo, E. 1930b/2010. On Boundary Numbers and Domains of Sets. In Zermelo (2010).