A Pragmatic Defense of Logical Pluralism

Cian Guilfoyle Chartier

A Pragmatic Defense of Logical Pluralism

Cian Guilfoyle Chartier

A Pragmatic Defense of Logical Pluralism

Cian Brendan Guilfoyle Chartier

A Pragmatic Defense of Logical Pluralism

ILLC Dissertation Series DS-2021-05

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

For further information about ILLC-publications, please contact

Institute for Logic, Language and Computation
Universiteit van Amsterdam
Science Park 107
1098 XG Amsterdam
phone: +31-20-525 6051
e-mail: illc@uva.nl
homepage: http://www.illc.uva.nl/

Printed and bound by Ipskamp Printing.

ISBN: 978-94-6421-303-4

A Pragmatic Defense of Logical Pluralism

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. K.I.J. Maex
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Online
op vrijdag 30 april 2021, te 10.00 uur

door Cian Brendan Guilfoyle Chartier
geboren te Montréal

Promotiecommissie

Promotores:	prof. dr. F.J.M.M. Veltman	Hunan University
	prof. dr. S.J.L. Smets	Universiteit van Amsterdam
Overige leden:	prof. dr. M.J.B. Stokhof	Universiteit van Amsterdam
	prof. dr. F. Berto	Universiteit van Amsterdam
	prof. dr. D. Cohnitz	Universiteit Utrecht
	prof. dr. J. Meheus	Universiteit Gent
	dr. M.D. Aloni	Universiteit van Amsterdam

Faculteit der Geesteswetenschappen

Contents

Acknowledgments

The journey to completing this thesis was long and not always certain, and the work owes its development to helpful remarks from too many people to list here. Primarily, Frank Veltman presided over the bulk of the thesis revisions, patiently, for many years, steering the project in the right direction. Sonja Smets vitally agreed to, along with Frank, see the thesis through to completion in the final year, Others who have read and usefully commented on parts of the thesis include Albert Visser and Johannes Marti. I am further grateful to the members of the doctoral committee: Martin Stokhof, Franz Berto, Daniel Cohnitz, Joke Meheus, and Maria Aloni.

Dora Achourioti and I co-hosted what would become the workshop "Truth Be Told Again"; she, among many other things, invited some of the best-known speakers and introduced me to the various philosophy mailing lists so it would be clear where the research that would become this thesis could be presented. Furthermore, Paul Dekker's LeGO Seminars served as an introduction to a number of novel perspectives.

Graeme A. Forbes at the University of Kent was pivotal in a change in focus from work on semantic theories of truth to work on the philosophy of logic; first with insightful discussions on the topics of realism and relativism in metaphysics, and second by allowing some of this material to be presented at the Evidence Seminar there. The work benefited from fruitful discussions at the Evidence Seminar with (not exhaustively): Daniel Auker-Howlett, David Corfield, Virginia Ghiara, Veli-Pekka Parkkinen, Gavin Thomson, Christian Wallmann, Michael Wilde, and Jon Williamson.

The work that would become the third chapter of the thesis was presented at several conferences (along with a LeGO Seminar and an Evidence Seminar), but particularly useful were comments from Graham Priest and Johannes Stern at the Munich Center for Mathematical Philosophy. The material that would appear in chapters one and two was presented on two occasions: first at the Evidence Seminar in Kent, and then at CLMPST 2019 in Prague. Critical remarks from David Corfield and Gavin Thomson were appreciated in the former talk, and from Andrea Sereni in the latter. Jody Azzouni's pointed criticism of logical pluralism, and a free copy of his book *Talking About Nothing*, were also much welcomed.

The common thread in all of this is the support of my wife, Alaina Schempp, without which this dissertation would not have been completed. We now welcome a new light in our lives, our daughter Mairéad Cora Guilfoyle Chartier Schempp.

Birmingham, United Kingdom   Cian Guilfoyle Chartier
March, 2021.

Chapter 0

Introduction

0.1 The Central Claim

The central claim of the thesis is that a logic is a formal presentation of a guide to undertaking a rational practice, a guide which itself is constituted by epistemic norms and their consequences. It is then maintained that there in general may be more than one "good" presentation, more than one "good" practice, and more than one way to conceive the practice. Call this *thoroughgoing logical pluralism.* The thesis consists in a defence of thoroughgoing logical pluralism, and a case for how taking such a perspective is helpful in addressing problems in logical revision, semantic paradoxes, and the incommensurability of logical theories.

It should be noted what kind of notion of logic we seek to define. Daniel Cohnitz and Luis Estrada-González in [17] helpfully outline certain distinctions that appear in various parts of the related literature. In particular, a logic as a *pure object* is a kind of mathematical structure also known as a *logical structure* such as a class of models or a proof calculus, or instances of such. A logic as a *non-pure object* is a phenomenon that has a logical structure, such as a valid argument form.[1] A logic as a *pure theory* is a study of some kind of logical structures, and accordingly a logic as a *non-pure theory* is a study of some kind of phenomena with logical structures.

We can apply these distinctions to our conception of logic, but it requires a little abuse of terminology. Here a *guide to undertaking a rational practice* may be a non-pure object if one presumes that such a guide has a logical structure, but

[1] Cohnitz and Estrada-González also follow Graham Priest's invocation of the notion of a "canonical application", which we go on to object to in chapter one, in order to distinguish certain forms of non-pure objects as "canonical" (e.g. argument forms) versus "non-canonical" (e.g. computer programs) according to the traditional use of logical structures. It is not our place to dictate future traditions.

we do not need this presumption. A logic in our desired sense, as a formal presentation of that guide in order to undertake that practice, has a logical structure that we impose in order to best approximate that practice. Logic so conceived is both an object and a theory: it is a pure object in that it is a logical structure of some kind, but it is also a non-pure theory of a guide to undertaking some particular practice in that it is normative for said practice. I said this required a little abuse of terminology because the guide having a logical structure is not necessary. Indeed it may transpire that the logic and the corresponding guide to practice diverge, but as we will argue in chapter two this may sometimes call for a revision of the practice rather than the logic.

A treatment of contemporary accounts of logical pluralism, starting with Beall and Restall's in [5] and [6], does not appear here but instead is saved for chapter one, in order to motivate and contrast with the account I provide in the thesis. Similarly, contemporary arguments for logical monism are saved for chapter one. Furthermore, the accounts of these works are themselves partial, with an eye to motivating the subject of the thesis rather than being intended as any sort of thorough exegesis of any of the particular works in question.

The second chapter engages (in broad disagreement) with a "anti-exceptionalist" strand in the recent literature on logical revision. An adequate summary of the anti-exceptionalist project is found in [35]. The third chapter requires a prior engagement with the literature on the semantic paradoxes, beginning with Kripke in [42], through Graham Priest in [52] and [54]. A broad summary of the problem the semantic paradoxes pose is to be found in [44]. More generally, a broad familiarity with deviant logics (paraconsistent and intuitionistic) along with elementary results in set theory and in models of arithmetic is assumed.

0.2 The Outline of the Thesis

The first chapter of the thesis outlines the aforementioned view of logic (demarcating what is meant by "formal" and "epistemic norms"), and defends it against the usual challenges to logical pluralism, notably that pluralism is self-defeating and Graham Priest's influential "collapse" objection that logical pluralism collapses into logical monism when considered. There is also a treatment of a few rival approaches: J. C. Beall and Greg Restall's logical pluralism ([5], [6]), Nikolaj Pedersen's logical pluralism adapted through Michael P. Lynch's account of alethic pluralism ([45], [51]), and the contemporary readers of Carnap's logical pluralism via linguistic pluralism ([15]), notably Stewart Shapiro and Teresa Kouri Kissel ([69], [40], [70], [41]). There is also consideration of the problem of incommensurability of logic: how it is that practitioners of two different logics understand one another. The conclusion is that practitioners understand each other through charitable interpretation, but charitable interpretation does not involve an assumption of an alignment of beliefs (as it often does in for instance [20], facing a challenge of solipsism), but instead involves constructing a formalisation that is our best fit of the epistemic norms we attribute to the other person such that, if she followed them, the norms would be of maximal respective utility with respect to our reading of her instrumental desires. This is the *practice-oriented principle of charity.*

The second chapter of the thesis defends the view that logical revision is a form of narrow reflective equilibrium with respect to the norms which constitute the practice that the logic is formalising, and the formalisation itself. The standard proponent of the view of logical revision being a matter of reflective equilibrium is taken to be W. V. O. Quine ([58]), but he was also a logical monist who was famously skeptical about the meaningfulness of deviant logic, understood in terms of *piecemeal* logical revision. Thoroughgoing logical pluralism and the practice-oriented principle of charity provide a basis in which the deviant logician is better understood. After all, the deviant logic is not understood as a piecemeal revision of another, but instead holistically as a formalisation of the deviant logician's practice, as constituted by the epistemic norms of maximal respective utility with respect to our reading of their instrumental desires.

A recent trend in the piecemeal approach to logical revision, itself partly in reaction to Quine, is seen in various approaches to revision under the umbrella of *logical anti-exceptionalism.* The name comes from the idea that there is nothing that distinguishes logical theories from theories of the natural sciences, in particular there is nothing especially a priori about the content of logical theories that distinguishes them from the a posteriori content of theories of the natural sciences. The logical anti-exceptionalists under study further argue that logic ought to be revised by means of *inference to the best explanation*, just as scientific theories presumably ought to be. Where they differ is what is meant by "best", and how they would face the challenge to classical logic of semantic paradoxes such as the liar. Timothy Williamson's idea of "best" leads him to rationalise sticking with classical logic ([93]). Graham Priest's idea of "best" leads us to revise classical logic ([57]). Ole Thomassen Hjortland's idea of "best" in [35] leads us to abandon universal quantification.

In the second chapter of the thesis the particular norms they use to justify the particular conclusions they reach are questioned. But logical anti-exceptionalism itself is more broadly attacked with a brief account of the normativity of logic, that distinguishes logical revision as a matter of merely *internal coherence* of epistemic norms and their formalisation (as opposed to revision of scientific theories). This is what is meant by narrow reflective equilibrium as opposed to *wide reflective equilibrium*, which is also discussed and rejected as an alternative.

The treatment of the anti-exceptionalists of the semantic paradoxes raises the question of how the form of logical revision explored allows us to reach conclusions about the suitability of various "solutions" and their upshot. Such a question is too broad to answer in its totality, as there may be many different starting points. But there is an interesting potential case study in how one accounts for *strengthened liar paradoxes* that appear to emerge in a recursive fashion whenever one tries to address the semantic paradoxes by means of distinguishing, in the meta-language, a set of paradoxical sentences (by Kripkean means, say).

In light of the above consideration, the third chapter outlines a family of pragmatic solutions to the liar paradox along the lines of the practitioner simply identifying the norms she wishes to keep, and changing her practice or formalisation in accordance with either keeping a truth predicate in her logical language or not, so either way paradoxical statements do not appear in the extension of a truth predicate in this language. This comes about through identifying the existence of paraconsistent (dual) Kripke fixed point-defined truth predicate extensions of a family of languages of arithmetic, starting from arithmetic and stretching into a hierarchy of predicates: the truth predicate, and a proper class of ordinal-indexed "paradoxicality" predicates. Such a language is defined and the existence of fixed points for the extensions of a hierarchy of predicates for paraconsistent models of arithmetic is proven.

Finally, the fourth chapter addresses in more depth an application of the practice-based principle of charity, and in particular how to address the *Problem of Communicability* that arises from asking what it is a classical formal mathematician can learn (about their own practice) from a nonclassical formal mathematician (in terms of their practice), and vice versa. In particular, the Problem of Communicability is to be accounted for on the basis of a characterisation of the *potential ability* or *strong knowledge-how* for the formal mathematician to make inferences. First the *Propositional Thesis* is ruled out: that the knowledge-how of a logician may be reduced to their knowledge-that of the propositions expressed by their logic's entailed sentences, or their knowledge-that of the propositions expressed by the sets of sentences that follow from others that they wish to accept as given. Intellectualist challenges to the primacy of knowledge-how by Jason Stanley and Timothy Williamson ([73]) are partially addressed with the aim of defeating the propositional thesis.

With the Propositional Thesis ruled out the remainder of the fourth chapter outlines how, through a modification of Craige Roberts' Questions Under Discussion framework ([62]) with the practice-oriented notion of charity differing from Kouri Kissell's own modification of the same ([41]), one can provide an account of how practitioners of different logics may understand and learn from each-other's proofs (to the extent that, given linguistic constraints, a formal translation sufficient to do this is possible). A few examples of the pragmatics of mathematical conversations are briefly worked through, allowing us to elucidate some cases of faultless disagreement coming out of trying to think of classical logic intuitionistically, or think of intuitionistic logic classically.

Chapter 1

On Logical Pluralism

1.1 Introduction

Over the last fifty years, work into all manner of nonclassical logics has become widespread if not predominant in the contemporary literature on logic. Yet among many philosophers of logic, there is only one correct logic, in a tradition going back thousands of years – the disagreement is instead over what that correct logic is.[1] The view that there is only one correct logic is *logical monism*. The view that there is more than one correct logic is *logical pluralism*.

What is meant here by "correct logic" is given by the pragmatic sense of some logic being the *right logic to use*, that which is not in general definable by the logic in question. This is not the familiar Tarskian notion of truth under which statements, or sets of statements, of a theory may be said to be true *in a model*, otherwise "logical pluralism" would just be vacuously true. The theory of all sentences of first-order predicate logic is that which is *true in all models* of first-order predicate logic, but that doesn't speak to its being a "correct logic" in the sense so intended. More to the point is the literature on the *normativity* of logic, concerned in part with how and whether logic can compel us to rightly change our beliefs, norms, or practices–this is one sense in which logic may be said to be *normative for thought*. There are other conceptions of the normativity of logic (see for instance the three conceptions of normativity in [74]) but for our purposes, it's good enough to say that if logic is normative for thought in the sense just given, then it is correct.[2]

[1]The question of who the first logical pluralist is is a matter of controversy, as the idea of what logic is changed considerably through the influence of Frege, Russell, and Tarski. Johan van Benthem in [82] traces logical pluralism's origins to Bernard Bolzano's *Theory of Science* in 1837, through the works of Charles Sanders Peirce beginning three decades later. It is argued in [59] that Hugh MacColl, a contemporary of Frege and Russell and an early pioneer of modal logic, might have been the first logical pluralist.

[2]Whether logic is normative for thought has itself been under contention; for instance Gilbert

Logical pluralism is the view defended here. The reason for this is that logic can serve more than one correct practice, this reason in turn being supported by much of the aforementioned last fifty years' work on nonclassical logic. Mathematicians and logicians are working towards different purposes and applications. As contemporary philosophy ought to serve as the fasteners keeping the sciences together and working in harmony, it is the role of philosophers of logic to help make these purposes explicit and possibly find new applications.

In the last 20 years, a number of different kinds of logical pluralism have appeared in the literature, and here a novel version, referred to here as *Thoroughgoing Logical Pluralism*, is introduced. Thoroughgoing Logical Pluralism, in two sentences, is as follows. Logic is a formal presentation of a guide to undertaking a rational practice, a guide which itself is constituted by epistemic norms and their consequences. Furthermore, it is a contention of Thoroughgoing Logical Pluralism that there are multiple correct logics when they are construed in this way.

By "thoroughgoing" it is meant that the sort of logical pluralism on offer is *comprehensive* in the following way: there may be multiple correct formalisms for multiple correct guides for multiple correct practices. In each of these cases of pluralism (multiple correct formalisms to a guide, multiple correct guides to a practice, and multiple correct practices) the differing formalisms/guides/practices may have their own epistemic value, which is to say they might each in an important sense be right to use.

Throughout this chapter and indeed the whole dissertation, there is an assumed focus on logics which are formal presentations of declaratives–rather than questions or interrogatives, as in inquisitive logics, or normative statements, as in deontic logics. The other forms of logical pluralism discussed here are typically (sometimes entirely) concerned with declaratives as well. Accounts of logics of declaratives should come first in an account of the role of logic in general, and an account of the role of inquisitive and deontic logics in the spirit of that of this dissertation can appear at a later date.

In what follows an account of Thoroughgoing Logical Pluralism will be outlined first, followed by a comparison of that account with some of the other proposals, while heading off some influential objections to logical pluralism.

Harman in [29] argues that formal logic is not a helpful guide to rational practice, and Gillian Russell in [64] maintains that no logic is normative in the sense just outlined. Contrary to these accounts, the second chapter includes a defence of the thesis that logics at least, perhaps at most, are normative for thought in that they revise our existing norms on the basis of narrow reflective equilibrium. But for now the normativity of logic is granted.

1.2 A Justification for Logical Pluralism

In this section what distinguishes Thoroughgoing Logical Pluralism from other accounts is established. Arguments in later sections will not only defend this account from principled objections to logical pluralism in general, but also contrast it with other accounts of logical pluralism in the extant literature.

We should, first, be clear that logic (in terms of its role) is in Thoroughgoing Logical Pluralism conceived as a formal presentation of a guide to undertaking a rational practice, a guide which itself is constituted by epistemic norms and their consequences. Other conceptions of logic will be discussed which will have nothing in common with this except for the apparently mundane view that logic is a "formal presentation", but even there we might still disagree on what we mean by "formal", which will receive further elaboration. Logical pluralism is advocated in three respects:

- There often may be more than one correct formal presentation of a "guide to undertaking a rational practice", with no one being obviously superior in correctness to the others, and the different formalisations do not even have to be extensionally equivalent. A relatively uncontentious example is the choice of ordinals in classical set theories; Von Neumann or Zermelo ordinals may be adopted. More contentious, but in turn more illustrative of an interesting pluralism, are families of supervaluational theories of truth, or revision theories of truth with different extensions depending on the initial conditions (for an overview see [8]).

- There may be more than one correct guide to undertaking a given rational practice. This accounts for the possibility of different families of epistemic norms constituting a guide to the same practice; the claim is also that in at least some cases both guides may be correct to use. There are, for instance, somewhat different justifications and motivations for intuitionistic mathematics; compare the BHK interpretation (briefly later in this section) with the "justificationist semantics" of assertion seen in [21].

- There is more than one correct practice. An example of how incompatible correct practices may be conceivable is worthwhile. You might find belief to reflect your willingness to take a bet, as Frank Ramsey did in [60], and come to take a Bayesian reading of belief. Or you might have a more cautious character and interpret belief intuitionistically, as Sergei Artemov and Tudor Protopopescu do in [2]. In the spirit of our pluralism, we can say that there may be (second-order) situations where it is worthwhile to bet before making a rigorous verification, and other situations where it is better to be more cautious. There may also be situations where it is not clear what approach is correct, but one practitioner or another has a personal preference for which one to take based on their own desires.

So in the above three respects, more than one logic may be correct. Guides to undertake a rational practice manifest in the practices themselves, and are constituted by epistemic norms.[3] What is compelling about justifying guides to rational practice rather than individual epistemic norms is that the guides, being manifested in practices that are productive or helpful in some way, justify themselves by their own value. The norms that underlie them might be either followed explicitly, or instead implicitly by simply following the practice. So whenever someone is engaged in a rational practice we can also say they have selected epistemic norms to follow, whether implicitly or explicitly.

Thus logic can be considered as a means to make epistemic norm-following explicit and, through formalisation, abstracted from extraneous details. This process comes about in the following stages:

- Through their genetics and past experiences, a practitioner has a set of *prior interests* that inform their choice of rational practice and *instrumental desires* that inform how they effectively undertake this practice. This is in keeping with Hilary Kornblith's contention in [39] that epistemic norms are informed by the "desires in a cognitive system that is effective at getting at the truth", but equivocates on whether the "truth" is something that rational practice aims at.[4]

- A practitioner's choice of rational practice, together with their desires, lead them to a selection of *epistemic norms*, hypothetical imperatives contingent on how well they fit their instrumental desires and prior interests, that constitute a guide for good practice. Some of the norms have more weight than the others in that they may be less amenable to revision.

- That practitioner *formalises* their epistemic norms into a logic.

- The logic entails consequences that contravene some of the epistemic norms the practitioner already committed to.

- Depending on the weights the practitioner has assigned the norms, the practitioner revises their epistemic norms with respect to their rational practice.

[3]Guides so conceived are referred to in [88] as *interpretations* which provide their own top-down justification for following the norms that constitute them. But we shall use "guide" instead, given the continued relevance of the Tarskian sense of the word "interpretation".

[4]There are at least two views to take: the first would be the view that the truth is a basis for epistemic evaluation, being something that rational practices collectively aim at (see indeed Kornblith in [39] but also [49]); the second would be that what we intrinsically value is the basis of epistemic evaluation, as in Stich's *The Fragmentation of Reason* ([75]). A discussion of what grounds epistemic value, or values, would be outside the scope of this thesis.

Inevitably due to the fallibility of practitioners put up against logical omniscience, the logic entails consequences that contravene some of the epistemic norms to which the practitioner is committed. So then depending on the weights the practitioner has assigned the norms, the practitioner revises their epistemic norms with respect to their rational practice. Narrow reflective equilibrium appears to be the most fitting characterisation of this process. An alternative approach to revision appears among defences of "logical anti-exceptionalism", i.e. the position that there is no clear distinction between logical and scientific theories (like the traditional distinction of logical theories being specifically a priori). This alternative is to regard revision as taking place in the form of an inference to the best explanation for contrary evidence. My criticism of this approach is the subject of the second chapter.

Some paradigm cases of differing practices are necessary to mention, along with norms that may be part of their guides. Already mentioned are the greatly different interpretations of belief of Frank Ramsey compared with Artemov and Protopopescu in [2]. But even more developed in the extant literature is the disagreement between the practices of intuitionistic and classical mathematics, corresponding respectively to families of intuitionistic and classical logics respectively.

Intuitionistic mathematics is partly constituted by at least two epistemic norms. The first norm we call the *proof-theoretic norm of meaning*: if you want to know the meaning of a proposition A, you must have a means to prove or refute A, and if A is a compound formula then its proof (refutation) is explained in terms of the proof (refutation) of the relevant constituent subformulas. The proof in question is provided recursively, as Anne Troelstra provides in [81]:

1. A proof of $A \wedge B$ is given by presenting a proof of A and a proof of B.
2. A proof of $A \vee B$ is given by presenting either a proof of A or a proof of B.
3. A proof of $A \to B$ is a construction that transforms a proof of A into a proof of B.
4. Absurdity $\bot$ ... has no proof; a proof of $\neg A$ is a construction which transforms any supposed proof of A into a proof of $\bot$.

Taken together, this is the BHK interpretation of intuitionistic logic. Michael Dummett and Dag Prawitz developed a different foundation for intuitionistic logic on the basis of a theory of meaning, fleshed out in the "justificationist semantics" seen in [21], but the above is in a similar spirit.

The second norm we call the *constructive norm of meaning* for mathematical assertions: if you want to know the meaning of a mathematical assertion, you must know the means of construction of all the mathematical entities that the assertion claims exist. Take for instance the account Heyting gives in the textbook [34] of first-order quantification for mathematical objects: we say that $\exists x : \rho(x)$ whenever a member a of a mathematical species S for which $\rho(a)$ is true has been constructed.

Both of these norms would run contrary to a *classical truth-conditional norm of meaning* for a proposition A, such as: if we want to know the meaning of a mathematical assertion, it is necessary and sufficient to know the conditions of the worlds under which the assertion is true. First, knowing the conditions of the worlds under which a mathematical assertion is true does not mean we have a proof of the proposition that the assertion expresses in the sense given by the proof-theoretic norm of meaning. Second, knowing the conditions of the worlds under which the assertion is true does not mean we know the means of construction of all the mathematical entities that the assertion claims exist. A classical truth-conditional norm of meaning might be part of a guide to a classical mathematical practice. It is beyond the scope of this thesis to argue the value of one particular practice or another, but it is to be shown how one can coherently think that both have value in each of their formal manifestations.

The idea of logic being underwritten by a (perhaps unarticulated) practice has a certain pedigree. L. E. J. Brouwer famously held that the faculty of mathematics came prior to logic in an account of mathematical practice. From what Brouwer called "the simple intuition of time" (see [12]) one understands the notions of recursion and continuity, and can thus engage in "the construction of intuitive mathematics". Logic for Brouwer is an application of these intuitions used to identify regularities in language, a "mathematical system of words" for which propositions such as axioms are but "linguistic images of connections" between basic notions. Thoroughgoing Logical Pluralism is a generalisation of this insight, that possibly heretofore unarticulated practices underwrite logic, but accounting as well for practices other than "the construction of intuitive mathematics", for instance the thought processes of classical mathematicians.[5]

Some of the later sections will include elaborations of: (1) how we cannot simply say the process of deciding between logics necessitates a collapse of logical pluralism into logical monism; (2) how having "prior interests" and "instrumental desires" does not lead us into methodological solipsism. A few words should be

[5] Moreover, Brouwer took a generally dim view of the relevance of logic to mathematics, at the height of the foundational crisis in mathematics. Logic, when divested of the expressive power to formulate vicious paradoxes such as Russell's, would be "no more than a defective expedient for men to communicate mathematics to each other and to aid their memory for mathematics". Contrary to this, the relevance of logic to revising one's own practice is crucial to the program of Thoroughgoing Logical Pluralism.

said about what is meant by "epistemic norms", "weight", and "formalisation" in the preceding story.

Here an instrumental view of what epistemic norms are, as in [39], is helpful to adopt. As an explanatory strategy, grounding norms in instrumental desires and prior interests has the advantage of evading the potential regress or circularity that comes from attempting to evaluate epistemic norms: if the evaluation were done with other epistemic norms, the question would be how those norms are evaluated; instead desires and interests provide a starting point.

Hartry Field in [26] invokes the idea of the norms that underpin logics having some weight ascribed to them that determines whether or not they should be revised, akin to being closer or further from the centre in Quine's web of belief. Thoroughgoing Logical Pluralism shares this idea and keeps to the spirit of his suggestion that "we don't choose an inductive method ab initio at all, we acquire one by genetics and training", as indeed the instrumental desires are acquired. But this limits our ability to treat of weights in any great detail, at least not within the confines of this thesis. Unlike with the weight of a norm, the *formalisation* of a guide to following rational practice is something which can be helpfully demarcated. In particular, John MacFarlane in chapter six of his PhD dissertation [46] provides a sufficiently general notion of formality in terms of categorical grammar.

Much space could be taken to try to defend each part of the account itself in more detail. But the more urgent issues are the principled attacks on logical pluralism, that we will find don't affect the present account. There will still be space to show the advantages of the present account against some similar rivals in the literature.

Thoroughgoing Logical Pluralism shares with Hartry Field the idea of norms having some weight that affect their ability to be revised, and keeps to the spirit of his suggestion that "we don't choose an inductive method ab initio at all, we acquire one by genetics and training", as indeed the instrumental desires are acquired. But this limits our ability to treat of weights in any great detail, at least not within the confines of this chapter. Unlike with the weight of a norm, the *formalisation* of a guide to following rational practice is something which can be helpfully demarcated. This is what is spoken of next.

1.3 Formalisation of Norms

Logic is conceived here as a *formal presentation* of epistemic norms. What is meant by formal is what John MacFarlane refers to in his PhD thesis [46] as "2-formality" in the following list of possible conceptions of the formality of logic:

1. **1-formality:** Logic is formal in that it provides the constitutive norms for

thought as such.

2. **2-formality:** Logic is formal in that it is indifferent to the particular identities of objects.

3. **3-formality:** Logic abstracts entirely from the semantic content of thought.

A central tenet of MacFarlane's PhD thesis is that the 2-formality of logic is not well-supported, and one might also have to assume the 1-formality of logic. So by this understanding (combining with the understanding of logic in the previous chapter) logic is a formalisation of a guide to a correct practice, *at the same time* providing the constitutive norms for thought as such. MacFarlane's argument presents a challenge to Thoroughgoing Logical Pluralism and deserves further consideration.

Before treating of how MacFarlane thinks we need the 1-formality of logic to support the 2-formality of logic, it should be emphasised that for a defence of Thoroughgoing Logical Pluralism, neither 1-formality nor 3-formality are acceptable. 3-formality requires abstracting away from such content as the manifestations of interests and instrumental desires, which are foundational to the present account. The 1-formality of logic turns out to be a central assumption of what I will treat of as the "collapse argument" against logical pluralism. In any case, if logic provided the constitutive norms of thought as such, there would be only one guide for good practice, that being given by the constitutive norms the logic provided; epistemic value pluralism would immediately collapse into epistemic value monism, undermining the need for any account of logical pluralism whatsoever.

In this section MacFarlane's argument for accepting 1-formality via his *Intrinsic Structure Principle* (henceforth ISP) shall be briefly treated, along with an explanation of just how MacFarlane assumes a role for logic in deciding questions in philosophy of language that turns out to be too strong. His argument is in some sense compelling, but the conclusion to draw from it is not 1-formality but rather a more modest relationship of logic to epistemic normativity suggested by Novaes in [22].

1.3.1 Intrinsic Structure Principle

In this subsection terminology shall first be set (from MacFarlane's thesis) for a *categorical grammar* which will be applied both here and later on. Later on, MacFarlane's three stages of semantic interpretation will be briefly outlined along with how, in terms of the categorical grammar, logic can be demarcated in terms of what MacFarlane calls "intrinsic structure". In what follows we see that we follow MacFarlane as far as his *intrinsic structure principle*, but depart as soon as he characterises that as a kind of 1-formality. While on his reading logic is

1-formal as well as 2-formal, the leap to 1-formality is one that we do not have to make.

Categorical grammar provides an account of logical syntax in terms of how *derived categories* of grammar are derived from *basic categories.* A basic category may be a set of basic syntactic entities such as terms and sentences. Any derived category is of the form (a, b) and is what combines with an expression of category a to form an expression of category b.

An expression is said to be *well-formed* if and only if either it is a lexical primitive, or it is the concatenation xy in the category b of two well-formed expressions x in the category (a, b) and y in the category a, or it is the function $\lambda x.y$ in the category (a, b) of x a variable in the category a and y a well-formed expression in the category b; n-place predicates are expressions that yield $(n-1)$-place predicates when applied.

As syntax corresponds to compositional syntax, so categories correspond to *types.* Corresponding to the basic categories T and S for terms and sentences are, respectively, the *basic types* O and V for the set of objects and the set of truth values. For each derived category (a, b) corresponds a type $(A \Rightarrow B)$ which is the set of functions sending the type A corresponding to category a to the type B corresponding to the category b. The correspondence can be given by an interpretation function i sending expressions and variables to truth values and objects in the familiar way.

Types like V may be said to have *intrinsic structure.* Consider (as in [1], this is also an example MacFarlane uses) a four-valued logic whose four values form an ordered lattice on V with respect to $\wedge$ and $\vee$ as meet and join operations respectively (and $\neg$ as the complement). With respect to the ordering $\leq$ on V, have t at the top, f at the bottom, $t > b > f$, $t > m > f$, and b and m incomparable with one another. The operations $\wedge$, $\vee$, and $\neg$ have truth value matrices invariant under the permutation switching b and n, leaving t and f unchanged, along with the identity permutation over V. These are the two permutations that preserve the ordering $\leq$ over V and so $\wedge$, $\vee$, and $\neg$ are considered formal in that they are indifferent to intrinsic-structure preserving maps of truth values *as well as* the identities of objects. Examples like these motivate talk of intrinsic structure before any general idea of what it is that makes orderings such as $\leq$ specifically *intrinsic.*[6]

[6]MacFarlane also appeals to a type of *moment-history pairs* in Belnap and Green's tense logic ([7]) as having intrinsic structure in the sense that V does. By contrast, he does not consider the modal accessibility relation in modal logics to, in general, have intrinsic structure. But the particulars of these arguments do not concern us here.

Before we go any further seeing where MacFarlane draws the line for what makes structure intrinsic, it is necessary to go into his general definition of formality. MacFarlane defines permutations of any type in terms of the same permutation over the basic types over which that type is defined. Given permutations σ^A, σ^B, σ^C... of basic types A, B, C... respectively, an *induced transformation* σ^Z on the type Z induced by σ^A, σ^B, σ^C... is defined inductively on the complexity of Z: if $Z = T$ where T is a basic type then for any $w \in Z$, $\sigma^Z(w) = \sigma^T(w)$; for any types X, Y, if $Z = (X \Rightarrow Y)$, then for any $w \in Z$, $\sigma^Z(w) = \sigma^Y \circ w \circ (\sigma^X)^{-1}$.

We then say that a semantic value w belonging to a type Z is *permutation invariant* if and only if for any intrinsic structure-preserving permutations σ^A, σ^B, σ^C... of types A, B, C... respectively, $\sigma^Z(w) = w$. In accordance with 2-formality, we can then say that w is logical, i.e. formal, if and only if it is permutation invariant with respect to intrinsic structure-preserving permutations.

So what we understand by formality so far amounts to considering what the logical and nonlogical constants are. All we are missing in the above demarcation of logicality is an account of just what kinds of structures are intrinsic with respect to the "intrinsic structure-preserving permutations". The idea is that what intrinsic structure there is should be seen at some point in the structure's interpretation. Key here is a distinction MacFarlane draws between three stages of interpretation, given as if in a chronological sequence:

1. The *presemantics* tell us what items in each type are assigned (which) semantic values.
2. The *semantics proper* tell us how the semantic values of expressions depend on the semantic values of their parts.
3. The *postsemantics* tell us how the semantic values of expressions relate to properties of their proper use.

The invariance criterion concerns the permutation of arbitrary items in a particular type, so it concerns the presemantics of a logic. MacFarlane's invocation of the *Intrinsic Structure Principle* provides a reason to demarcate logic at the postsemantic level:

> **Intrinsic Structure Principle:** One should use a type with just as much intrinsic structure as is needed for the postsemantics, and no more.

Up until this point, MacFarlane's demarcation project is in broad terms compatible with the present one, as a guide to what is "formal" about a "formal presentation of a rational practice". But MacFarlane sees the Intrinsic Structure Principle as a species of 1-formality–which is not acceptable to us:

> Sensitivity to *intrinsic* structure does not compromise the general applicability of a logical notion, because intrinsic structure belongs to a type in virtue of the most general purpose of logical theory: the study of the semantic relationships that hold between sentences solely in virtue of their capacity for being asserted and used in inferences. On this ground, one might say that notions that are sensitive only to intrinsic structure are applicable to *thought as such*, independent of its particular subject matter, and that the laws governing these notions are norms for thought as such.

It is one thing to say that formal logic is sensitive to (and only to) intrinsic structure, and that intrinsic structure is only that needed for the postsemantics (i.e. "for being asserted and used in inferences"), but this doesn't lead us to 1-formality without a significant leap. The leap is in assuming there is one particular kind of "thought as such" to which the logical notions are all applied. On the contrary, intrinsic structure belongs to a type not "in virtue of" a single "thought as such", but to the particular practice that the logic is meant to serve. In pithy terms, there is not one kind of thought in this sense but many.

We have already demarcated practice as something constituted by epistemic norms, which logic is meant to "formalise". This formalisation is a rendering into a language with logical and nonlogical operators, which are adequately characterised by a modified notion of 2-formality in that logical operators are permutation invariant with respect to intrinsic structure-preserving permutations. Intrinsic structure is provided by the demands of the postsemantics, i.e. minimal conditions given by epistemic norms for the logical language to be an appropriate rendering of the practice as expressed.

Still relevant, though, is the question of the sense in which logic is a normative discipline, i.e. that the formalisation tells us something significant about what is right, or what we should do. That will be the subject of the remainder of this section.

1.3.2 On the Normativity of Logic: Logical Descriptivism

A crucial question for the logical pluralist is just what sense a logic is said to be *right*. This turns on the extent to which logic can be a *normative* enterprise, given the view of logic in question.

A natural place to start is with the view that logic isn't normative at all, which we call *logical descriptivism*. Gillian Russell in [64] provides a carefully-articulated disentanglement of logic from normativity. A logical theory has normative consequences, but only when coupled with normative assumptions. Otherwise, statements of the theory are "normatively inert" in that they constitute a merely *descriptive* account of what is true and false with respect to that theory. More generally, logic provides a descriptive account of the laws of truth, in a Fregean sense, which for Russell is "analogous to the descriptive psychology of happiness, hunger, lying, and rule-breaking". Assuming Hume's Law, logic isn't normative.

As Russell herself maintains this claim has relevance for a defence of her own view of "logical pluralism", but it is helpful to focus here on the descriptivist aspect of her account. Descriptivism is a convenient view in that it completely dodges the "collapse objection" against pluralism that is mentioned later. Her views are tied into a broader anti-exceptionalist project, that there is nothing distinguishing logical theories from scientific and mathematical theories. Setting aside for the moment various issues with anti-exceptionalism, which will appear in the next chapter, it will nonetheless be beneficial to focus instead on the claim Russell makes that logic isn't normative.

First, note that Russell takes for granted that scientific and mathematical theories aren't normative:

> Newtonian mechanics tells us that certain things are true, e.g. $F = ma$ (force is mass × acceleration) and hence we can derive from it the conditional fact that if $F = 24N$, then $ma = 24N$ too. This means that at least one member of $\{F = 24N, ma \neq 24N\}$ is not true. Putting it together with ["You ought not to believe a sentence is true when it is not"] allows us to draw the conclusion that one ought not to believe that every member of $\{F = 24N, ma \neq 24N\}$ is true. So in the presence of norms that connect truth to things one ought to do, physics has normative consequences. But I think we would and should say that physics isn't really normative; the 'oughts' in these conclusions have their source in widespread norms connecting belief and truth. Physics merely tells us which things are true.

This implicitly sets a number of challenges for what can be considered a normative discipline. As Russell acknowledges, there is a weak sense in which Newtonian mechanics is obviously normative: under certain circumstances, we ought to believe that if $F = 24N$, then $ma = 24N$. Her claim is that logic, which she conceives in a Fregean sense to be a description of the laws of truth, is only a normative discipline in that sense.[7]

[7]Gottlob Frege himself maintained that logic is a normative discipline, but it is not clear whether he intended to convey that logic is normative in the same sense that ethics is; [48] includes a relevant discussion.

Unlike Russell logic is here taken to be a normative discipline, and in a stronger sense than Newtonian mechanics, but a review of the concerns over how to argue the discipline is explicitly normative may be necessary.

1.3.3 On the Normativity of Logic: Bridge Principles

In the ensuing section, we respond to Russell's challenge by outlining a "bridge principle" which ensures that logic is in some sense right, or tells us what we should do with respect to following or revising our practice. Before then, we outline what is meant by a "bridge principle" and what sorts of challenges there have been to formulating one. The implementation of the principle will be elaborated on in the second chapter.

To respond to Russell's challenge, John MacFarlane's work on normativity is relevant. In [48] he argues that in order for logic to be considered a normative discipline, there should be rules connecting a normative "ought" to either truth or validity. Moreover, these rules should be at least partly constitutive of the notion of validity. In [47] MacFarlane presents many potential examples of such rules, which he calls *bridge principles*, connecting a normative "ought" to either truth or validity. It is instructive to view an example:

> (O) If $P_1, P_2...P_n \vDash Q$, then one ought to see to it that if they believe $P_1, P_2...P_n$, then they also believe Q.

This bridge principle (O) is ruled out in Gilbert Harman's influential arguments in [29] that logic does not play a significant role in rational practice. In particular, the problem of logical omniscience presents itself: it is not plausible that someone believes all the logical consequences of what one believes. It is also, Harman argues, not desirable: "One should not clutter one's mind with trivialities." This is often referred to as the *excessive demands* objection.

The *paradox of the preface* presents another challenge. Say you have just written a book on astrophysics. You ought to individually believe every individual statement you've made in the book, but being fallible, you also ought to believe the negation of their conjunction. Yet as long as we have closure under conjunction, (O) tells us we ought to believe the conjunction.

MacFarlane suggests an alternative bridge principle that can account for both criticisms:

> (R) If $P_1, P_2...P_n \vDash Q$, then one has reason to see to it that if they believe $P_1, P_2...P_n$, then they also believe Q.

Yet as MacFarlane acknowledges, (R) is too weak. Just because one has reason to also believe Q doesn't mean that is what one ought to do. In principle it may be that the entailment relation never compels us to believe Q, even though it gives us reason to see to it that we do. Then the question arises of why logic tells us anything about what we ought to believe. As Catarina Dutilh Novaes acknowledges in [24], the desires for bridge principles that are on the one hand sufficiently strong and on the other hand able to accommodate Harman's challenges might be in conflict.

A helpful approach Florian Steinberger has in [74] to deal with the apparent dilemma this sets is to distinguish between different kinds of norm:

- An *evaluative* standard: a third-person principle that evaluates the conduct of an omniscient subject.
- A *directive*: second-person commands to a non-omniscient subject;
- An *appraisal*: third-person grounds to attribute praise or blame to a non-omniscient subject.

For our purposes, evaluative standards are not helpful to use as bridge principles, as the intent is to defend a view of logic as being manifested by our practice as non-omniscient subjects. This leaves directives and appraisals. A practitioner might follow directives to make use of logic. Someone assessing the competence of a practitioner might follow directives for logic to tell them how the practitioner should act.

Apart from directives being second-person and appraisals being third-person, the two kinds of norms might be substantively different in their demands. For instance, one way of weakening (O) that MacFarlane rejects is (O'):

> (O') If one recognises that $P_1, P_2 ... P_n \vDash Q$, then they ought to see to it that if they believe $P_1, P_2 ... P_n$, then they also believe Q.

MacFarlane rejects (O') because of what he refers to as the *priority question*: that a bridge principle ought to apply to someone even when they don't know it. The norms should apply even in our state of ignorance. Yet the priority question does not concern second-person directives, as it is unreasonable to expect people to act on directives they have not received. The priority question may still apply to appraisals.

It is useful here to take stock. Bridge principles should apply even in our own state of ignorance (the "priority question"), and one should not be able to derive too many trivial demands from them. When bridge principles are appraisals, we struggle with excessive demands. When bridge principles are directives, they don't apply in our own state of ignorance when they should. This leaves us at an impasse.

Thankfully, there is already an account of normativity that combines directives and appraisals in a manner that sidesteps both difficulties. Precedent here can be found in Robert Brandom's project of inferential semantics (in [10] and [11]), where beliefs are understood as doxastic commitments in a social practice of giving and asking for reasons. Indeed logic for Brandom "supplies the expressive resources needed to make explicit–to put in judgeable form–the semantic and pragmatic bases of judgment". We pass over Brandom's project of inferential semantics as a whole, but there is a parallel with the more modest aim to treat logic as a formalisation of epistemic norms that serve as the basis for judgment, that in terms may provide cause for further reflection on our practice. Novaes already follows this spirit in [24] by making bridge principles with respect not to what an individual ought to believe, but instead with respect to what two opposing agents in a dialogue, which Novaes calls Proponent and Opponent, ought to put forward and grant:

- (P1) If $P_1, P_2 ... P_n \vDash Q$, then Opponent ought to see to it that, if she has granted $P_1, P_2 ... P_n$ and Proponent has put forward Q, then she grants Q.
- (P2) If $P_1, P_2 ... P_n \vDash Q$, then if Opponent has granted $P_1, P_2 ... P_n$, then Proponent may put forward Q.

Here (P2) in some sense works as a directive for Proponent as it meets the excessive demands objection. (P1) on the other hand satisfies the priority question as an appraisal of Opponent. On the surface, the Proponent can take the role of practitioner and Opponent can take the role of scorekeeper.

On this account a few objections are apparent. First, the demands are insufficiently strict on Proponent: presumably there are many Q's that a human Proponent may put forward but would not. Moreover, the demands on Opponent are still too strong, at least if Opponent is supposed to be a human adversary. Novaes suggests that an individual working on a logic problem may act out the roles of both Proponent and Opponent, but then it is not clear what someone is being asked to do: what would that individual be committed to? Finally, for our purposes, it is not clear what role logic has to play in informing the already competent practitioner.

What is proposed, instead, is to consider Opponent to be not a fallible person but the manifestation of the formalisation of the practitioner's choice of logical consequence. The practitioner might recognise that for some class of models I, $P_1, P_2 ... P_n \vDash_I Q$. Then the following directive gives a normative role to validity:

> If you recognise a set of sentences Σ such that for each $Q \in \Sigma$, $P_1, P_2...P_n \vDash_l Q$, and you grant $P_1, P_2...P_n$, then you ought to see it to it that either every Q is granted, or that some of the norms invoked in granting $P_1, P_2...P_n$ or defining $\vDash_l$ are revised.

Here the practitioner is bounded under obligation to either grant every Q, or else revise the norms that were already assumed. The revision that is made would depend on the weights the practitioner has already assigned to their norms, but the choice to grant each Q is theirs. One rationally takes the path of least resistance: first remove those norms with the least weight for which it suffices that the resultant logic would not have (some formulation of) the undesirable Q's as a consequence, while at the same time assigning new norms and weights in line with the method of reflective equilibrium in [28]. The case for this approach will be argued in more detail in the next chapter.

The cost of the practitioner not accepting every Q would be high: it would be up for debate whether they would still even be engaged in the same practice, but this would be a question of what determines a particular practice rather than what role logic is meant to play in it.

1.4 Challenges to Logical Pluralism

Much of the following shall consist of treatments of principled objections to logical pluralism, with an eye on Thoroughgoing Logical Pluralism. First, briefly, a few typical formulations of monism are mentioned (and scrutinised), followed by a series of typical objections to logical pluralism helpfully outlined by Graham Priest. Of particular importance is the popular *collapse objection* to logical pluralism, where it is shown how in light of the practice-oriented nature of Thoroughgoing Logical Pluralism, the collapse argument ultimately fails. It is then shown how the usual challenge of conceptual relativism does not apply.

1.4.1 Monism

Logical monism has a long-standing tradition in classical philosophy. This has come from a conventional view of logic being intertwined with what we now think of as metaphysics, going from Aristotle through Kant and Hegel. The discipline of logic has changed radically since the mid-19th century to the point of being a more disconnected study of different logics, many of which are nonclassical in some respect—constructive logics, paraconsistent and paracomplete logics, substructural logics, non-monotonic logics. The traditional conception of logic might explain why despite this, monism still is endorsed by most working philosophers.

Nevertheless it is admittedly hard to square monism with demands for logical revision, and what is called "classical logic" has indeed been significantly revised in the last 150 years. So the arguments for monism made by the monist proponents of logical revision, Timothy Williamson and Graham Priest, shall be scrutinised.

Williamson's logical monism could be seen as implicit in what he takes to be the close relationship between logic and metaphysics.[8] As he views logic as indifferent to context and methodology, with a universal reading of the quantifiers, there is barely any room left for pluralism, apart from a pluralism about some minor terminological variations which he concedes in [92] as a "boring truth" and a formalism about logic which he rejects. In this role there is (modulo boring truths) one true logic, motivated by his abductive criteria in [91].

There is a strange aspect of the approach: if the logical principles are supposed to be universal in scope, they should presumably have a bearing on Williamson's abductive methodology as well. For instance, one can say, "Either the law of excluded middle holding unrestrictedly is a good fit with the evidence, or it is not." Excluded middle would be imposed through an implicit assumption. One might then try to argue that it would be absurd not to unrestrictedly accept the law of excluded middle, or that we should accept it because we haven't found enough cases where it doesn't hold. Where we can go with the abductive methodology is heavily influenced by our starting point, and Williamson appears (in [92]) to have accepted this, but the question of what starting point to choose remains mysterious.

So much for Williamson's motivation for monism. The boundaries of Graham Priest's monism are clearly delineated. Priest refers to the formalist view of logics as the *many pure logics* view, which he takes to be "uncontentiously correct" while rejecting formalism about logic as a whole. Like Williamson, Priest also makes reference to the terminological view of pluralism as being true but uninteresting. Again the major issue arises from taking logic to be normative for thought.

Priest speaks of logic being normative for thought in a strong sense, as being evaluative of the right way to reason *about* particular situations. In particular he speaks of the study of the right way to reason as the *canonical application* of logic to reasoning about situations. Priest calls *theoretical pluralism*, which he disputes, a pluralism of theories about the canonical application relativised to particular situations.

[8]For instance, in [91], logic has the role of supplying "a central structural core to scientific theories, including metaphysical theories, in essence no more above dispute than any other part of those theories", in line with the anti-exceptionalist project.

Priest's view of how logics become theories in this sense unfolds in the following way. According to Priest, people essentially reason in the vernacular. Logics provide a concept of validity in a formal language. We want an account of validity in order to establish what holds in actual or hypothetical situations given what we assume about them. Here debates can emerge essentially about the applicability of logical laws in a domain or with respect to a set of situations, and from those debates a stance that some of these debates could be avoided by an embrace of some kind of theoretical pluralism. Yet this approach Priest objects to in principle, for reasons given in the next subsection.

1.4.2 Priest's Four Kinds of Attack

Priest has made several responses to Beall and Restall's pluralism (as seen in [5]), but many of them have been helpfully summarised in [56]. This can serve as a summary of challenges to logical pluralism as a whole, which deserve a response.

Priest's attacks on this view can be grouped into at least four kinds:

1. *(Against relativism)* Allowing for the meaning of logical connectives to shift with the context and situation leads to an incoherent view of what logic is.

2. *(Expressive limitations)* Certain kinds of pluralism, it seems, cannot honestly be expressed. If a classical logician, so the argument goes, advocates the use of intuitionistic reasoning in certain situations, they are not being intuitionists. After all, some intuitionistic mathematics is not consistent with classical mathematics.

3. *(Collapse arguments)* The kind of pluralism that different kind of objects have different properties, or different logical laws hold about them, can be expressed in terms of one overarching logic.

4. *(Realism)* Validity is a real thing that exists independently of our best theory of what it is. One may apply formal logics to different ends, not for what Priest calls the *canonical application*.

A considerable amount of the objections to logical pluralism since Beall and Restall's proposal have been concerned with the collapse argument, which itself often hinges on a realism about validity. Tthe collapse objection shall now be responded to in detail. The first two claims are accounted for later on with a specifically *practice-oriented* formulation of the interpretive principle of charity.

1.4.3 Responding to the Collapse Objection

The essence of the forthcoming response to the collapse argument is that what is meant by "collapse" is practice-relative and relative to the starting point one chooses. First it will be helpful to see just what collapse means, in terms of a metalinguistic array of functions on either deduction systems or classes of models that determine the most suitable of these for a given problem. This array can be as complex as the practitioner wishes, but there are already examples in the literature of how the choice of collapse function cannot be neutral.

Priest's formulation of the collapse argument is framed against Beall and Restall's defense of a notion of validity defined as "truth-preservation over different classes of situations".[9] Priest argues that validity could only be truth-preservation over all situations, and an argument for this is also an argument against the current approach. This is where the collapse argument finally appears:

> Let s be some situation about which we are reasoning; suppose that s is in different classes of situations, K_1 and K_2. Should one use the notion of validity appropriate for K_1 or for K_2? We cannot give the answer 'both' here. Take some inference that is valid in K_1 but not K_2, $\alpha \vdash \beta$, and suppose that we know (or assume) α holds in s; are we, or are we not entitled to accept that β does? Either we are or we are not: there can be no pluralism about this. In fact, the answer is that we are. Since s is in K_1, and the inference is truth-preserving in all situations in K_1. In other words, if we know that a situation about which we are reasoning is in class K, we are justified in reasoning with validity defined over the restricted class of situations K.[10]

Given $\vdash_{K_\alpha}$ a *particular notion of validity* defined over a class of situations K_α indexed by I, with K the class of all situations, and T_α the set of theorems of $\vdash_{K_\alpha}$, define a *collapse function* $f : \Pi_{\alpha \in I} T_\alpha \to T$ to be a function sending the set of theorems of every particular notion of validity to a set of theorems T of a *general notion of validity* $\vdash_K$ defined over K. The result, for Priest, is some notion of validity that can accommodate dialetheia, for those cases where different classes of situations conflict on the truth of propositions.

The peculiar feature of Priest's particular formulation of the collapse argument is the acceptance of propositions accepted as valid in a given class of situations. If $\alpha \vdash \beta$ is valid in K_1 and $\alpha \vdash \neg\beta$ is valid in K_2, then by Priest's lights we are entitled to accept both β and $\neg\beta$ given α. It is easy to imagine alternatives, such

[9] A critical treatment of Beall and Restall's approach appears in a later section.

[10] Despite the use of the symbol $\vdash$ instead of $\vDash$, Priest is referring to a model-theoretic conception of validity, essentially as truth-preservation over all classes of situation.

as a classical supervaluationist approach where β is always rejected and $\neg\beta$ always accepted whenever $\alpha \vdash \beta$ is not accepted with respect to some class of situations. Varying the collapse argument with different "one true logic"s as a conclusion is left as an exercise to the reader. Needless to say, the collapse arguments may be responded to all in one go.

This is the challenge that such an argument presents against the current approach: if logic is practice-relative, then it is also situation-relative, given that a practice p may be adopted in a class of situations K for which validity accords to that particular practice. Given the above argument, if the classes of situations K_1, K_2, K_3... corresponding to practices p_1, p_2, p_3... have any situations in common, then they have equivalent notions of validity. There are two possibilities: either there is one notion of validity for all situations up to equivalence, or there are classes of situations corresponding to rational practices that are completely disjoint with respect to the notion of validity that they manifest. Priest rejects the second possibility, as he believes there is a "core of universally correct inferences" in that some of these core inferences, such as conjunction elimination, are valid in every situation. This leaves us with the first possibility.

The relevant issue, in general, is that Priest's starting point is that we are reasoning about a situation, but typically this does not happen. Without getting into the particulars of the relations that might hold between situations, a situation so conceived is a set of propositions or a *partial* world, more generally a *possibility*. Conventionally, model-theoretic validity is defined in terms of truth in a class of possibilities, and there is good reason for this. For, contrary to Priest, the question of whether ψ follows from ϕ is not whether ψ follows from ϕ in some given situation as such. The question is rather what possibilities are compatible with ψ following from ϕ. It might transpire that the remaining possibilities radically conflict with our purposes, so logical revision takes place by reflecting on this.

The motivation for this emphasis is to avoid the cart coming before the horse. We reason in order to advance the practical knowledge of a community of inquirers, where there is a general understanding that we don't know what the possibilities are. So we do not start with an explicit possibility and reason about what follows inside that possibility. Moreover, we should in principle be able to change our mind about Priest's *starting point* situation s being a possibility at all, not just about what follows from s being a member of certain classes of situation.

The choice of situation s to reason about is not the only problem with the collapse argument; its very formulation involves making some metalogical assumptions. For instance, the collapse function is provided in terms of a metatheory defining both logics. Yet there are numerous examples showing that metatheory cannot be neutral, see for instance the discussion in chapter seven of [69]. Among the examples given, a particularly pertinent one is this: one could express intuitionistic sentential logic in a classical metatheory; but one could also express Brouwerian counterexamples in said classical metatheory, but not in an intuitionistic metatheory. Each Brouwerian counterexample is taken to be a *possible theorem* of intuitionistic sentential logic with respect to the classical metatheory, but not with respect to the intuitionistic metatheory. This results in a different input to the collapse function in each case, and in general a different output.

The lesson to take from this is that since the choice of metatheory is ultimately founded on practical grounds, and the choice of metatheory is seen to inform how the collapse function is defined, the so-called collapse of pluralism into monism is itself relative to one's practice and also one's own starting point. However, there is another path to take: perhaps the collapse argument still works as long as there is one best practice among them (regardless of our best theory of what it could be) which corresponds to the canonical notion of validity that Priest appeals to. That is to say, monists face the demand that in arguing for the correct logic, we must appeal to realism about higher powers such as "the" correct way to reason. Indeed, the class of valid relationships between truth-bearers is for Priest "determined by the class of situations involved in truth-preservation, quite independently of our theory of the matter". Priest calls this the *canonical application* of logic, or applications of what may be referred to as the *canonical notion of validity.*

If the canonical notion of validity existed yet no practitioner could ever rationally believe they had it or were approaching it, the "one best practice" claim wouldn't hold up in the collapse objection against logical pluralism. After all, then there would be no rational means to decide the collapse function. So the "one best practice" claim must also come equipped with a claim that we are compelled to find this canonical notion of validity at the rational end of inquiry. One might then have faith that pluralism would collapse into the One True Logic so long as we are being compelled to find the canonical notion of validity. Call this position *end-of-inquiry monism.*

One need not take a position on whether end-of-inquiry monism is true. Even if it were true, it would make little difference, as in the face of quaddition-esque sceptical challenges about the meaning of $\vdash_K$ we would have no way to know for certain that some inference is that of *the* notion of validity at the rational end of inquiry. And for the collapse argument to then work we would need that level of certainty, as otherwise we would not be able to isolate the relevant classes of situations with respect to which the collapse function could be defined. Merely formalising a best fit of our rational practice does not require that kind of certainty.

So much for the collapse argument. In the next section we should wish to distinguish Thoroughgoing Logical Pluralism from the methodological solipsism that may emerge from what I call a *representational account* of metalogic.

1.4.4 Metalogic as Conceptual Scheme or as Manifestation of Practice

In the preceding sections we have seen a pragmatic case for Thoroughgoing Logical Pluralism, and a refutation of the frequently-pitched argument that logical pluralism collapses into monism. Different practitioners use different logics for different purposes. There are at least two possible accounts of what is meant by "using different logics": the first being that the practitioners are working in a metalogic whose theory serves as a reflection of a conceptual scheme for their work, the *representational account* of metalogic, the second being that the practitioners are engaged in different practices for which their respective metalogics serve as a formalisation, the *practical account* of metalogic. In this section two reasons shall be given for why the second account is to be preferred.

The first reason why we should prefer the practical account is that metalogics themselves should have a justification for being adopted, but as the theory of the metalogic is a reflection of our conceptual scheme, the justification of the theory must be part of the theory itself, and thus viciously circular. This is a similar issue as that raised by Lewis Carroll in "What the Tortoise Said to Achilles" ([16]); justifications of the use of modus ponens required the use of modus ponens, leading to an infinite regress. Ultimately the regress was solved by Bertrand Russell in [63] with the notion of a *rule of inference* outside of the theory itself, which modus ponens is typically taken to be.

My alternative proposal is that in the order of explanation, a scheme of epistemic norms comes before the metalogic itself, and one can be a pluralist about which norms to accept. To accept the norms is simply to be engaged in a practice that the norms are intended to serve: all that is involved in a change of these norms is a change in strategy, either implicit or explicit. Being a pluralist about epistemic value, one can manifest a different metalogic serving a different practice towards a different value.

The second reason why we should prefer the practical account is that the representational account leaves us with no room for explanation for how substantive disagreements can take place, as practitioners working in different conceptual schemes cannot justify their different choices to one another or relate to the choices the others have made. Timothy Williamson's criticism of logical pluralism in [92] becomes relevant here, which can be summarised as follows: if we assume the practical role of a logical connective as how one thinks one ought to reason with it, or how one actually does reason with it, our resulting account of inference fails to account for the mistakes we are likely to make. So we must see the role of connectives as how we (individually) ought to reason with them. But all parties in a debate about logical connectives must take on the responsibility for establishing their meaning, and reach a consensus, unless the debate were to collapse into a form of total incommensurability where faithful agreement could never be made. So all parties in a debate ought to have the same meaning for the connectives, collapsing pluralism into agreement as Williamson suggests:

> Many deviant logicians make it clear that they are not stipulating idiosyncratic new meanings for the connectives but rather are proposing new theories about the old meanings of those connectives, according to which their logical powers have been mischaracterized by classical logic. Normatively, they are holding themselves responsible to the meanings of the connectives in a public language, employing the connectives with those publicly available meanings in order to engage in unequivocal public debate with defenders of orthodoxy... since all parties to the debate are to be held responsible to the public meaning of [some connective, the inferential role specifying how the subject should reason with that connective] is the same for all of them.

As an argument for logical monism this is much too fast. The meanings of the connectives that a practitioner is prepared to be responsible for may be just what was at issue in the first place; the stipulation of a public language for each pair of practitioners is too strong. Furthermore, it is clear from experience of attending conferences on logic that even without a public language, typical practitioners in a dispute about differing logics can disagree on the meaning of the connectives, and thus their practical roles, but still understand one another. The question is how to account for this. Without a public language requirement, how is mutual understanding between practitioners possible?[11]

The principle of charity can serve as a way of reconciling mutual understanding with substantive disagreement about the practical role of the logical connectives, with the simple thought from Donald Davidson in [20] that people already have enough of a solid basis of agreement on many matters that they can still communicate. To be clear, Davidson's principle of charity is that we make the most sense of others by optimising agreement.

---

[11] My objection to the requirement of a public language, and the forthcoming response, also applies to Jody Azzouni's "external discourse demand" in [4].

Williamson in [90] attacks Davidson's principle of charity, but defends his own: we make the most sense of others by maximising what we take others to know. Where we know the meaning of the logical connectives, we by default expect others to know the same meaning. Here we can see how logical pluralism must on Williamson's account collapse into either monism or total incommensurability. Indeed, especially as Williamson takes knowledge to be propositional knowledge or knowledge-that, this form of interpretive charity is excessively strong. For instance, it requires us to accept that for every set of propositions Γ that an interlocutor knows, they also know every proposition entailed from Γ.

Still contentious is Davidson's contention that what we and our interlocutors agree on are the propositions expressed by the sentences in our respective languages; if we've already interpreted these propositions, we've made a judgment on their truth, so "optimising agreement" would just mean superimposing our judgments over those of our interlocutor's. Obviously we're supposed to interpret the propositions as our interlocutor does, but how do we have access to this interpretation? We do it by having a handle on our interlocutor's practice, and supposing how we should interpret these propositions if we were to act as our interlocutor does.

We have arrived at *the practice-oriented principle of charity*: we make the best sense of others when we suppose they are following epistemic norms with maximal epistemic utility with respect to our possible interpretations of what their instrumental desires could be.

Whenever we formalise the same practice in different ways, or approach the same practice in substantively different ways, there is a source of disagreement. We assume our interlocutor is acting in good faith and trying to follow epistemic norms that best suits their instrumental desires. Our understanding of their instrumental desires comes prior to our understanding of their interpretation of the propositions in question, and such an understanding is possible because most of our instrumental desires are shared.

This concludes the responses to criticisms of logical pluralism. In the following section we shall move on to accounts of other extant logical pluralisms, which are either more limited or do not hold enough water against the collapse or relativism objections.

1.5 On Some Other Pluralisms

The discussion of pluralisms shall start with a brief treatment of Field's pluralism and Varzi's, followed by Beall and Restall's pluralism, criticism of which leads us to two other kinds somewhat closer in spirit: first Lynch and Pedersen's pluralism about the concept of truth motivating logical pluralism, and then Shapiro and Kouri Kissel's project of logical instrumentalism (motivated by a revival of Carnap's).

1.5.1 Hartry Field on Logical Pluralism

Hartry Field in [26] espouses a view of logical pluralism that, like the present account, is founded on a pluralism about epistemic normativity. In earlier sections we saw some parallels to mark some aspects in which this account agrees with Thoroughgoing Logical Pluralism. Now it would be helpful for us to focus more on the differences, which essentially boil down to Field also holding the view that logical theorems should be evaluated on whether they make true commitments about what follows from what *tout court*. In this way, practitioners that end up with different theorems in the same language are engaged in a factual disagreement. What Field allows for are differing epistemic norms manifesting themselves as different rules of inference.

It is important, in this context, to see where Field is coming from: though he is an anti-realist about epistemic normativity, and a fictionalist about mathematical objects, he is a realist about scientific theories, logic grouped among them. His account of "relativist expressivism" is a precursor to the current account, but how epistemic norms impinge on logic differs in two notable respects:

1. Though logical implication is thought of as being normative, "ground level connectives" that are used to compose sentences of the logic itself are, by Field's lights, not normative as the sentences are merely factual claims.
2. Field has in mind proposals for logic as an *all-purpose method*. Local goals aren't the aim here, but rather a global guide to thought as such, at least with an aim to establish and preserve truths.

One might wonder what sort of pluralism is left after these concessions are made: the answer is that two logics with the same language might both be regarded as good relative to different epistemic priorities for how they work through their consequences rather than what those consequences are, i.e. the two logics might have the same set of logical consequences but a differing definition of logical consequence itself, or differing rules of inference.

Field's account of pluralism differs from Thoroughgoing Logical Pluralism in its attitude to the role of epistemic goals. We should not regard logic as providing global norms for thought or reason, but rather as a manifestation of the norms of a particular practice. It may be that two logics with the same language have been conceived with different enough practices in mind that their differences in theory do not amount to a straightforward difference in fact. At best, it may reflect a difference in attitude.

1.5.2 Pluralism About Logical Demarcation

Alfred Tarski in [78] and more recently Achille Varzi in [83] have advocated a scepticism about logical demarcation which has manifested itself as a form of logical pluralism. The scepticism can be stated as follows: there is no clear distinction between logical and nonlogical constants.

In various parts of the logical pluralism literature, Tarski and Varzi's pluralism has been viewed as perhaps acceptable but not very interesting, in Field's sense. On the contrary, Varzi makes clear that his scepticism about logical demarcation should lead to a radically relativist position, more in line with Carnap's. The remainder of this section focuses on Varzi's treatment in [83].

Varzi's account is a denial of a very widely-held intuition of the role of logical constants in the semantic notion of logical consequence. For Varzi, "all bits of language get their meaning fixed in the same way, namely, by choosing some class of models as the only admissible ones"–*including* the supposed "logical constants", which keeps them from being logical constants in the first place.[12] He then provides an account in terms of types filtered through categorical grammars of how this could be. As our earlier discussion about just these types sheds some light on the discussion, it's worth looking at this account in more detail.

Consider that we have built a language with a categorical grammar including the category S of sentences. Let's additionally say that we are interpreting this in a two-valued semantics $2 = \{0, 1\}$ where 0 and 1 are the ordinals representing falsehood and truth respectively. Unary sentential operators are of the category (S, S), and binary sentential operators are of the type $(S, (S, S))$. We can define classical negation and addition without recourse to Tarski's truth definitions; Varzi does this as follows:

- Let s_i be in (S, S). If $s_i = \neg$, then $i(x) = 1 - x$ for all $x \in 2$;
- Let s_j be in $(S, (S, S))$. If $s_j = \wedge$, then $i(x)(y) = x \cap y$ for all $x, y \in 2$.

So "logical constants" can be defined without recourse to Tarski's truth definitions. There is, for Varzi, nothing *intrinsic* in the notion of a model that requires

[12] This may be a questionably uniform treatment of reference, many names ("orange") may be obtained by demonstration, while others ("Évariste Galois") are historical and others ("Galois field") arise from conventions about things or practices. But in any case, allowing for variation of choice of logical constants leads to a corresponding variation in the choice of logical models.

logical constants to have the same definition, and indeed there may be models for which they have substantively different meanings.

The above view is broadly in step with Thoroughgoing Logical Pluralism, but it's still not logical pluralism in a particularly controversial sense. After all, there is no account of how we come to the right choice of logical constants, or why we should start by varying the choice of logical constants in particular (as opposed to, for instance, the definition of entailment). There's a thought along the Carnapian relativist lines that there ultimately are no good answers to the question of the correct logic to use. It shall be argued in a discussion of Kouri Kissel's pluralism that this is probably not the right way to go, but before that will first treat of a popular account of pluralism that very much focuses on the definition of entailment rather than the logical constants.

1.5.3 Beall and Restall

J. C. Beall and Greg Restall in [5] and [6] outline and defend a notion of logical pluralism which is purportedly divorced from epistemic and metaphysical concerns. They do not speak of "one true logic", though they appear to agree with the tenets of classical logic ahead of paraconsistent and intuitionistic logics. Instead, they define what *a logic* is from a standard definition of validity that they purport to understand as the *pretheoretic notion of logical consequence*, and they are pluralists in that as there are multiple instances of this definition of validity, there are multiple logics. In the recent literature on logical pluralism this formulation has predominated; the attacks on logical pluralism that we have seen are mostly responding to Beall and Restall.

Their general definition of validity is (quoting from [6]) the following:

1.5.1. Definition. An argument is said to be valid$_\chi$ if and only if, in every case$_\chi$ in which the premises are true, so is the conclusion.

A logical monist would either deny the above general definition of validity, or maintain that there is only one such valid$_\chi$. Beall and Restall believe there is more than one. There are certain ranges of case$_\chi$'s in which certain arguments are said to be valid$_\chi$, and certain cases in which they might not be. In particular, Beall and Restall define an *admissible instance* of their definition of validity to be one obtained by a specification of the case$_\chi$'s, and of the relation *is true in a case*, such that the instance satisfies the "settled role of consequence" and makes necessary, normative, and formal judgments about validity.

A few words should be given to what Beall and Restall mean by admissible instances of their definition of validity. Validity is a relation, most generally conceived between collections of propositions. Logical consequence is the "fundamental topic" in logic, which in turn "concerns itself with the evaluation of arguments". The judgments about validity are said to be necessary in that the truth of the premises of a valid$_\chi$ argument necessitate the truth of the conclusion with respect to all the case$_\chi$'s; valid$_\chi$ arguments are necessarily truth-preserving. Beall and Restall agree that judgments about validity are merely normative insofar as people should accept the conclusion of a valid$_\chi$ argument if they accept every case$_\chi$ in which the premises are true. As a norm, it may be trumped by other norms.

So, to summarise: a logic is an admissible instance of Beall and Restall's definition of validity. Such an admissible instance must provide its validity as a norm of rationality, though one which may be trumped by other norms. An admissible instance either provides the constitutive norms for thought as such, is indifferent to the particular identities of objects, or abstracts entirely from the semantic content of thought. Most particularly, an admissible instance is necessarily truth preserving.

Among these admissible instances of their definition of validity they explicitly mention classical logic, relevance logic, intuitionistic logic, and paraconsistent logic, and gesture at many others to the point where the only valid$_\chi$ argument they all have in common is that a conclusion A follows from the single premise A. This commonality, however, is suggestive of how pluralistic Beall and Restall are willing to be. Their notion of logic necessitates a view of logical consequence that is reflexive, transitive, and monotonic, which rules out a wide range of substructural logics as being "logics" in their sense. This is a significant limitation of their approach which they acknowledge.

There is one other limitation of their approach worth mentioning. Logics which count as admissible instances are, in their account, normative for thought even if not mutually coherent: "if an argument is valid, then you somehow go *wrong* if you accept the premises but reject the conclusion" (their emphasis). But there is no principled justification for one logic trumping another for a particular purpose, though it may for particular cases:

> The pluralist claim is that, given a body of informal reasoning (that is, reasoning not produced in a particular system of logic), you can use different consequence relations in order to analyse the reasoning. As to *which* relation we wish our own reasoning to be evaluated by, we are happy to say: any and all (admissible) ones! Our arguments might be valid by some and invalid by others, good in some senses and bad in others. But that is not the end of the story. Once we learn that our argument is bad in some sense–for example that a verification of some

> premises will not itself be a verification of the conclusion–we will not necessarily *thereby* reject the usefulness of the argument. It depends, of course, on whether the given kind of verification preservation is important to the task at hand. Similarly, one might say 'Look, here is a Tarskian model of first-order logic which makes your premises true but conclusion false.' What should we do in response? Well, it depends. ([6], their emphasis)

Different logics may be normative for thought but not in an interesting sort of way on their own, only in light of importance to the "task at hand". But there is no general principle underlying what this could be. Thoroughgoing Logical Pluralism starts from differing values and practices, prior to formalisation and logical evaluation. The values and practices make clear the "task at hand" before evaluation takes place.

1.5.4 Logical Pluralism from Pluralistic Concepts

To be an *alethic pluralist* is to be a pluralist about what truth is.[13] There are ways in which one can be an alethic pluralist while not also a logical pluralist, and vice versa. For instance, one can hold that there are multiple concepts of truth, yet each concept provides comprehensive support for their favourite logic. Conversely, one can be a logical pluralist without being a pluralist about truth: Beall and Restall's pluralism is essentially this, as their conception of logic is based on truth preservation, and while validity and cases are defined relative to a class, "true" is not.[14]

In light of the above, one might wonder what the relevance of alethic pluralism is to logical pluralism. Yet a story of how logic is normative is elegantly grounded in a story of how truth is also normative. This can be seen from a *functionalist* view of truth, as Michael Lynch does in [45], which is to say it can be seen by viewing truth in terms of its discursive role as a rigidly-designated term that nonetheless *manifests* itself in different properties that are domain-specific. In what follows we shall elaborate on how this comes to be, but then remark on a criticism of Lynch's approach that suggests that truth does not have to play this functional role, and the insights taken from Lynch's functionalist approach can

[13] In what follows, we shall assume that alethic pluralists are pluralists about the concept of truth, rather than just pluralists about what truth is in natural language, though often the two conceptions of alethic pluralism turn out to be at least close relatives, as advocates such as Michael Lynch speak of finding the "ordinary" or "folk" concept(s) of truth.

[14] In [6] Beall and Restall distinguish truth simpliciter from *logical truth*: "a claim A is logically true if and only if, for some class X of cases determining a logic, A is true in every case in X". Note that logical truth here is not defined in terms of truth simpliciter, but in terms of truth-in-a-case.

be viewed in a more general pluralism about epistemic normativity.

As I've mentioned, for Lynch truth is meant to be viewed in terms of its discursive role, or its *truth-role.* To be more specific, "the properties that can determine that propositions are true are those that play the truth-role." Moreover, a sentence is true if and only if the sentence has a property that plays the truth-role. Certain *core truisms* are intended to be partially constitutive of any property P of objects x that plays the truth role, including:

- *Objectivity:* [given an arbitrary proposition p] "The belief that p is true if, and only if, with respect to the belief that p, things are as they are believed to be."
- *Norm of Belief:* [given an arbitrary proposition p] "It is prima facie correct to believe that p if and only if the proposition that p is true."
- *End of Inquiry:* "Other things being equal, true beliefs are a worthy goal of inquiry."

The core truisms are intended to provide a non-exhaustive list of features of distinct properties that play the truth-role, each conceived as different *manifestations* of truth. This is the sense in which Lynch's view of truth is pluralistic. Properties that play the truth role are, apart from core truisms, otherwise filled out by more substantial features that fit the domain conditions. For example, Lynch gives an example of a property of truth as that of *superwarrant* according to a coherentist view of warranting:

> [given an arbitrary proposition p] *Superwarrant:* The belief that p is superwarranted if and only if the belief that p is warranted without defeat at some stage of inquiry and would remain so at every successive stage of inquiry.

With all this in mind, let's look at how all this is relevant to logical pluralism. Lynch acknowledges that one does not need to be a logical pluralist in order to be a functionalist about truth, but also sees pluralism as a natural position to take in light of this. Why is that? A natural way for logical pluralism to emerge through truth functionalism in is through viewing truth-in-a-domain as the same thing as what Beall and Restall call truth-in-a-case, that being truth in a model, then isolating certain classes of cases as paradigmatic of some of the "more substantive" truth-role features, such as superwarrant. So truth is not typically expressed in the object languages of the logics in question, but in the metalanguage where the truth-role is some essential features of the notion of truth at a model. This is just the case that Nikolaj Pedersen shows in [51], arguing (by an appeal to Beall and Restall's truth-in-a-case) that the feature of superwarrant is paradigmatic of intuitionistic logics, and a truth property with certain "classical features" is paradigmatic of classical logics.[15]

We saw earlier how a story of how logic is normative can be grounded in a story of how truth is also normative. And we can see how that is: Norm of Belief and End of Inquiry provide truth with a normative character, and a definition of validity in terms of truth would accordingly have normative force. Additional norms for truth-role properties that are characteristic of the domain of inquiry would provide partial grounds for distinguishing between different logics. The situation, however, becomes considerably more complicated when considering attempts to interpret *compound statements*, i.e. statements where "true" is used in more than one distinct sense, where the different senses are not mutually compatible.

Lynch invokes a principle of "modesty" as a rule to account for these cases: "where a compound proposition or inference contains propositions from distinct domains, the default governing logic is that comprised by the intersection of the domain-specific logics in play." Yet even if read charitably, the intersection is often considerably weaker than either logic in question.

The interpretation of compound statements is a much harder problem than this thesis can treat in depth, but logical pluralism does not require us to treat these cases, particularly if (as in the conception of logic undertaken with Thoroughgoing Logical Pluralism) the rules come first, the analysis of truth coming later. The problem here is that if Lynch starts from a conceptual analysis of truth as a basis for logical pluralism, then he must account for these compound cases. For while we can question whether people are following epistemic norms that we have set prior to our formalism, and prohibit them from reasoning in ways that unjustfiably deviate from practice, there is nothing on Lynch's account to stop people using the word "true" in two different ways.

The issue is not just with Lynch's bridge principles; the upshot here is that the grounding of logical pluralism on pluralism about the extension of a particular concept used to ground the normativity of logic may not be pluralist enough. For example, something similar would happen if we try to ground logical pluralism on a pluralistic conceptual analysis of intentional action (such as Sosa's account of knowledge as apt judgment in [71]) or of knowledge or belief. The lesson is that what grounds the normativity of logic must come prior to the logic's expression, which is the case for us and also in the pluralism of Shapiro and Kouri Kissel, to

[15] The "classical features" in that case are that every state of affairs pertaining to reality either obtains or does not obtain, and for any arbitrary proposition p, $\neg p$ is true if and only if p is not true.

which we now turn.

1.5.5 Logical Instrumentalism

Another take on logical pluralism is a revival of Rudolf Carnap's project of rational reconstruction in [15], largely seen in the work of Stewart Shapiro and his student Teresa Kouri Kissel. This can be grouped under the heading of "logical instrumentalism", the title of Kouri Kissel's PhD thesis [40].

Shapiro in his book *Varieties of Logic*, ([69], hence VL) endorsed a wide-ranging view of logical pluralism: that the choice of logic is context-sensitive and interest-relative in regard to the mathematical structures being considered. Kouri Kissel in her PhD thesis, and Kouri Kissel and Shapiro in [70], advance this program.

Shapiro in VL endorses the *Hilbertian perspective* about mathematics as applied to logic, summarised by the following:

> Consistency, or some mathematical explication thereof, like satisfiability in set theory, is the only formal criterion for legitimacy – for existence if you will. One might dismiss a proposed area of mathematical study as uninteresting, or inelegant, or unfruitful, but if it is consistent, or satisfiable, then there is no further metaphysical hoop the proposed theory must jump through before being legitimate mathematics.

What this doesn't tell us is what is the right logic to use for our purposes. How do we know which logical connectives we want to use, and which interpretation of them we want to have? How do we know the inferences we make are the ones we are supposed to make? In VL this is an open subject:

> The present account, which extends the Hilbertian approach to logic, also owes an account of how mathematical theories, with different logics, can be applied in the scientific study of non-mathematical reality. And it should determine the proper logic for such applications. I won't attempt anything like that in detail here, mostly because the subtleties of application would take us too far afield. In general, it would be a matter of keeping track of which inferences are valid for which subject matter, mathematical or otherwise. And, of course, the various applications might mix together, illuminating each other. We'd have to figure out which logic to use in these mixed contexts.

This is exactly the problem one might hope to begin to address. Fortunately, this has partially been undertaken by Shapiro's student Kouri Kissel, who in her

PhD thesis [40] makes a case for the "instrumental" role of logic that undertakes the greater part of this problem. Very briefly, Kouri Kissel's instrumentalism can be summarised as follows. In the rational practice of a discipline like mathematics, one observes certain norms. The role of logic is to serve as a reconstruction of these norms.

Yet Kouri Kissel is wary of prescribing logic for those who are already competent practitioners, which Carnap himself was inclined to do in [15]. Moreover, the meaning of logical terms is given strictly inside the context of discussion, the question of the meaning *simpliciter* of logical terms being an *external question* outside the realm of inquiry (another evocation of Carnap). Kouri Kissel in [41] provides a way to formalise piecewise disagreement about the meaning of terms, without providing an overarching answer to such external questions as what "the" right way to reason is, through an application of Craige Roberts' *question under discussion* discourse framework (henceforth QUD) in [62]. The details of this application are discussed in more detail in the fourth chapter.

Attempts for logicians to legislate to the sciences require the invocation of evaluative judgments rising from the comparison between logics, but the questions that arise of what one logic is good for relative to another are Carnapian external questions.[16] Thus the Questions Under Discussion framework, as Kouri Kissel uses it, provides a basis for understanding lexical ambiguity for an already highly regimented form of discussion, where substantial disagreement over the value of the different logic used cannot be characterised.

Indeed, in Kouri Kissel's thesis she argues that logic ought not to legislate to "properly functioning sciences", among which would include physics, linguistics, and mathematics. As Kouri Kissel says, "Legislating to these sciences comes down to telling the science they are wrong about their practices on philosophical grounds. I take it that as a point of humility, we ought not to do this." This point of humility is central to her ensuing case against logical monism. On the contrary, a reasonable hope is that logic *can* serve as a legislator to the sciences, but it would be a lot easier with a pluralistic approach.

What logic does not do on Kouri Kissel's reading, yet in some cases it demonstrably can do, is provide grounds for its own revision to better fit norms of rational practice, and provide grounds to practitioners for second thoughts – see the earlier remarks on logic and normativity. There is a sense in which both of these purposes can readily be seen: logics provide a framework for computational verification that can outrun the mental capacities of their practitioners. Manifold consequences beyond the easy reach of the practitioner can be verified computa-

[16]Thanks to Peter Railton for relevant remarks on Carnap.

tionally. This includes some mathematical results, take for instance the proof of the four color theorem using the Coq proof assistant ([27]).

1.6 Conclusion

The aim of this chapter has been to provide some sort of justification for a great deal of work on nonclassical logics, for that work stands in disquieting contrast with reflections on the normativity of logic that generally support one correct way of reasoning, often (but not always) some form of classical logic. The justification given here is pragmatic in a manner similar to Wittgenstein's: the practitioner first has (consciously chosen or unconscious) rules that govern their practice in a given domain, then forms a logic to formalise these rules. When faced with unforeseen consequences, they judge whether the consequences demand a revision of some of these rules.

Details have not been provided about how the rules are to be revised, upon a judgment that a revision is to be made rather than an acceptance of the consequences. Broadly speaking, as already mentioned, a revised system of norms should be devised along a path of least resistance and through the method of reflective equilibrium. The details of this shall be filled in in the following chapter, but only in part, as values (with which weights are assigned) are practice-relative.

A central distinction of Thoroughgoing Logical Pluralism is that it is founded on a pluralism about epistemic value. Different epistemic norms serve different purposes, and may be manifested in different logics for which there is no ground for disagreement on substance. The priority of these norms and their plural manifestation have been seen to undermine the collapse argument against logical pluralism. Faultless disagreement and substantive disagreement are both possible and understood through the practice-oriented principle of charity.

In contrast to the instrumentalism of Teresa Kouri Kissel and Carnap's methodological solipsism, it appears desirable to embed the role of logic in an "external" epistemic theory, complete with a bridge principle tying logical consequence to epistemic normativity. The motivation for this is that contrary to the idea that the logician must be humble in the face of the practitioner they merely emulate, it is reasonably clear that logic outruns the capacity of the theorist to think through consequences, and this is ultimately a big part of why logic is useful.

There has so far been no comment on the distinction between logic and mathematics, as mathematical pluralism has its own literature.[17] Where discussions

[17] A general introduction can be found in [38]

on mathematical pluralism touch on the acceptance of both classical and intuitionistic mathematics the prescriptions are the same. A pragmatic conception of the role and content of mathematics might be developed along the lines of this thesis, but would require a considerable amount of new detail. For instance, there has not been the space to discuss how mathematical practice relates to scientific practice as a whole, or what bridge principles might compel us to try to prove a certain theorem or apply mathematics in a particular way. Such a study would be fruitful given the depth and variety of contemporary mathematics, which is often not understood by practitioners working in slightly different areas of research.

The means by which a domain of practice is established, and the rules revised, could be seen as a basis for the project of *logical revision*: justifying a change in stance of the right logic to use in a given domain of practice. There has been a growing interest in the project of logical revision in general, but by means of an *inference to the best explanation* of the right kind of reasoning. In the following chapter some objections to this idea are raised, and the opportunity is taken to present, in more detail, just how logical revision could work in the current approach.

Chapter 2

On Logical Revision

Since the increased interest in logical foundations of mathematics in the 19th century, logical paradoxes provided persistent sources of anxiety for philosophers and logicians about the legitimacy of parts of mathematical practice – real analysis, for example. Paradoxes such as Lewis Carroll's dialogue in "What the Tortoise Said to Achilles" ([16]), and limitative theorems such as Kurt Gödel's incompleteness theorems and Alfred Tarski's undefinability theorem of truth, provided motivation for mathematical logicians to carefully distinguish between (respectively) the axioms and rules of inference of a logical theory, and a logical theory's object language and metalanguage.

Eventually the anxieties about the legitimacy of mathematical practice eased. Zermelo-Fraenkel set theory with the axiom of choice (ZFC) came to be accepted as a standard foundation of the vast majority of mathematics, including classical model theory. Classical first-order predicate logic became widely regarded among most mathematicians as the one true logic.[1] But dilemmas raised by logical paradoxes persisted, and papers on the semantic paradoxes continue to be produced at a startling rate.

Especially since Saul Kripke's seminal "Outline of a Theory of Truth" and Graham Priest's provocative advocacy of paraconsistent logic, debates about the semantic paradoxes began to centre around whether or not to revise the rules of classical logic. With that in mind it is surprising to see that sketches of a general theory of *revision* for logical theories have appeared relatively recently. Graham Priest and Timothy Williamson have recent papers each outlining a similar abductive methodology for logical revision: but the former concludes that paraconsistent logic prevails over classical logic, the latter concludes the opposite. In response to both, Ole Thomassen Hjortland outlines his own abductive

[1]Work on intuitionistic mathematics continued, but was divorced from Brouwer's goal to provide an alternative foundation of mathematics.

methodology to support a pluralistic view of logics.

The position that all three writers set themselves against is Quine's skeptical attitude to logical revision, that a "change in logic" amounts to "a change in subject". In response they all embrace in some form a project of *anti-exceptionalism*, that logical theories are not based on analytic or a priori truth and are continuous with scientific theories, and thus should be revised by the same kind of abductive methodology.

By an "abductive methodology" we mean that old claims are revised in favour of new claims, on the basis of the new claim being the simplest explanation that best fits the current evidence. In the cases we are discussing, the "explanation" is a product of a logical theory, and the "evidence" is a phenomenon of reason one might hope to capture with that theory. In this paper, it will be argued that Williamson's commitment to "logical strength" is needlessly restrictive, and in the case of Hjortland and Priest there is reason to believe they are not truly being anti-exceptional. It will be argued that their problems have a similar root, which all anti-exceptionalists share: they have misconceived the relationship of logic to scientific theories.

The next section is a note on Quine's skeptical attitude to logical revision, which provides different grounds for the rejection of such an attitude than those of the anti-exceptionalists. The various anti-exceptionalist approaches will then be treated: ultimately it shall be explained how Priest, Williamson, and Hjortland reach such different conclusions on the basis of an apparently similar methodology because they start out with different conceptions about "best fit". Following this, there are objections to Priest, and Hjortland, and then to Williamson's emphasis of theoretical strength. It shall be argued contrary to Williamson that his anti-exceptionalist treatment obscures the manifold applications of logic. Instead an approach to logical revision is outlined which sets the goal of logical revision to be narrow reflective equilibrium with respect to one's logic and norms of practice.

2.1 Quine on Metalogic

The project of anti-exceptionalism is opposed to the viewpoint, expressed most famously by Quine, that piecewise logical revision is not possible. Quine's case shall be briefly argued along with a response to it along the lines of the previous chapter on logical pluralism, which will then be a backdrop for the discussion of anti-exceptionalism that follows.

Quine's case in [58] is succinctly put in the following:

> To turn to a popular extravaganza, what if someone were to reject the

> law of non-contradication and so accept an occasional sentence and its negation both as true? An answer one hears is that this would vitiate all science. Any conjunction of the form ‘ $p.\neg p$ ’ logically implies every sentence whatever; therefore acceptance of one sentence and its negation as true would commit us to accepting every sentence as true, and thus forfeiting all distinction between true and false.
> In answer to this answer, one hears that such a full-width trivialization could perhaps be staved off by making compensatory adjustments to block this indiscriminate deducibility of all sentences from an inconsistency. Perhaps, it is suggested, we can so rig our new logic that it will isolate its contradictions and contain them.
> My view of this dialogue is that neither party knows what he is talking about. They think they are talking about negation, ‘ $\neg$ ’, ‘not’; but surely the notation ceased to be recognizable as negation when they took to regarding some conjunctions of the form ‘ $p.\neg p$ ’ as true, and stopped regarding such sentences as implying all others. Here, evidently, is the deviant logician’s predicament: when he tries to deny the doctrine he only changes the subject.

The way to make sense of Quine’s remarks is in light of his *holistic* view of logical theories, constituted as a whole of a set of logical truths, which themselves constitute the meaning of the logical connectives. Piecemeal revision of logic is not possible in principle because any allegedly piecemeal change in logic involves a change in meaning of the logical connectives. Advocates of piecemeal revision confuse a difference in meaning with a difference in “mere” implication, and are simply talking past the classical logician.

This viewpoint has been criticised from a number of perspectives. Some early commentators, such as Adam Morton in [50], argued that revisions of logic that do not contradict any statement of classical logic do not constitute a change in subject. This is to say that Quine’s thesis may be true for some piecemeal revisions of logic, but not others. Others, such as Juliette Kennedy in [37], have argued in terms of set theoretic metalogical foundations that certain concepts such as *computable function* and Gödel’s *constructive set* are “formalism free” in that they do not undergo a significant shift in meaning even with a shift in underlying logic (so long as they can be expressed at all). This is to say that Quine’s thesis may be true with respect to some concepts of investigation, but not others. Another more radical view is taken by the anti-exceptionalists we see here: that the merits of logical revision are not to be evaluated by the original logic’s own lights (and its reading of the logical connectives) but by abductive inference to the best explanation. So piecemeal logical revision can in principle be merited, and the paraconsistent logician Quine alludes to may do this by appeal to some abductive methodology.

The response taken here, instead, is as follows. Logics are formalisations of some rational practice. There are at least two ways in which logics can differ in substance: they may be different kinds of formalisation of the same practice, or formalisations of essentially different practices. The most firmly-developed understanding of a *difference in practice* being manifested in a difference in logic is that between classical and intuitionistic mathematicians.

How is a difference in practice to be understood? By the *practice-based principle of charity*: we make the best sense of others when we suppose they are following epistemic norms with maximal epistemic utility with respect to our possible interpretations of what their instrumental desires could be. Charity allows us to distinguish when a proposal for a piecemeal change in logic amounts to a change in subject, as we establish the epistemic norms that would best fit the proposed change of logic and determine whether they are radically different from our own. If we find that the best fit with the logic is a system of epistemic norms that is in agreement with our interests and instrumental desires, we evaluate on the basis of said interests and desires whether it would be a more suitable logic for our purposes.

This approach is open to one obvious objection: that practitioners do not have the cognitive resources to interpret each other charitably in this way. This failure might happen in one of two ways: first, one practitioner might fail to see a difference in practice where there is one; second, one practitioner might interpret the other as engaged in a different practice when there was no need to. The way to respond to this objection is to see that failures in either direction are either open to correction (where a better interpretation can be formulated in the subject's language), or indicative of expressive limitations of the subject's language. It is only when we hit those limitations, and have no imminent desire to revise our language, that the current picture fits Quine's "change of subject" remark.

Our perceptions of epistemic utility emerge from our own practices and goals. It is, in any case, a duty of the practitioner to see when another is engaged in a deviant version of their own practice or a different practice altogether; charity does not add any difficulty to this already extant picture. Charity merely serves as part of an explanation of how piecewise logical revision is possible in terms of this duty.

In the next few sections, various accounts of establishing the relevant interests and instrumental desires for evaluating logic by means of some particular abductive methodology will be treated; it will be argued that these accounts are not suitable, and ultimately no particular abductive methodology is suitable.

2.2 Timothy Williamson's Case for Classical Logic

Timothy Williamson provides a case against the "neutrality" of metalogic in [92], and provides an abductive basis for comparing and revising logical theories in [93] that strengthens the case for classical logic. Taken together, they provide the case that the semantic paradoxes do not provide a basis for revising logical theories, and we ought to reach that conclusion in a classical metalogic. We summarise Williamson's arguments and show how it presents a challenge for logical pluralism.

In adopting an abductive methodology for the revision of logical theories, one first examines evidence of a logical phenomenon that requires explanation. Then one finds the logical theory that provides the simplest explanation and best fit of the evidence. For Williamson, what distinguishes a logical theory as fitting the evidence is that it "verifies some of its predictions as well as falsifying none of them". The measure of best fit scales up with the amount of predictions verified. Thus if T and $T*$ are two logical theories and T is logically stronger than $T*$, then T is at least as good a fit with the evidence. Moreover, there are differing kinds of competing theories, which can be privileged based on how *fundamental* they are. For Williamson, metalinguistics is less fundamental than logic. So faced with a choice to revise a metalinguistic theory that would explain a metalogical phenomenon for a logical theory that explains logical phenomena, the logical theory is privileged.

Viewing "best fit" in the way Williamson does, it is particularly difficult to provide a good case to revise classical logic, due to the strength of classical theories. Take the semantic paradoxes. Semantic paradoxes are metalinguistic phenomena, and a disquotational theory of truth from which semantic paradoxes may be inferred is also metalinguistic. Revising classical logic, by (say) restricting the application of modus ponens, would weaken the logical theory to accommodate a metalinguistic theory, which for Williamson is an unacceptable compromise. Thus the disquotational theory of truth is sacrificed to preserve classical logic: a theory of truth may not have the T schema hold everywhere, for instance.

Williamson views the semantic paradoxes as a relatively clear threat to classical logic. He is much more brief about the competing claims of classical and intuitionistic mathematics. Assuming that classical and intuitionistic propositional logic are being considered in "a single already interpreted language", classical propositional logic is to be preferred, because $p \vee \neg p$ is a theorem of classical propositional logic but not intuitionistic propositional logic, and every theorem of intuitionistic propositional logic (under such a formulation) is a formulation of classical propositional logic.

There is a strange aspect of the approach: if the logical principles are supposed to be universal in scope, they should presumably have a bearing on the abductive methodology as well. For instance, one can say “Either the law of excluded middle holding unrestrictedly is a good fit with the evidence, or it is not.” After all, one would have already assumed we were reasoning classically. One might then try to argue that it would be absurd not to unrestrictedly accept the law of excluded middle, or that we should accept it because we haven’t found enough cases where it doesn’t hold. Where we can go with the abductive methodology is heavily influenced by our starting point, and Williamson appears (in [92]) to have accepted this, but the question of what starting point to choose remains mysterious.

Even if we take for granted that Williamson’s abductive methodology in principle leaves logic open to revision, in practice his criteria for strength combined with “best fit” strongly privileges classical logic.

2.3 Priest’s Model of Theory Change

Graham Priest is a proponent of paraconsistent logic, but he also advocates the universality of logic and logical monism. For him the semantic paradoxes motivate not just the use of a paraconsistent logic, but the *revision* of logic. Rather than deciding between logics in what adequately describes a phenomenon, he advocates changes in logic in order to accommodate all phenomena. In [57] he provides a sketch of a model of logical revision: the *Weighted Aggregate Model* (henceforth WAM). As Priest himself suggests, use of WAM is compatible with logical pluralism. We briefly compare it with Timothy Williamson’s abductive methodology.

WAM is a program of theory choice, where for Priest a *theory* is a form of explanation which “provides an account of the behaviour of certain notions (some of which are non-observational) and their interconnections... but also provides an explanation of these facts... [and] its acceptability can be determined only by some sort of process involving evidence and argument.” This notion is meant to appear general. It is meant to encompass *scientific theories*, such as Newtonian mechanics, along with *logical theories*, which in general are accounts of validity matched with a corresponding theory of meaning.

Say at least two interlocutors disagree on a single theory among m choices, $T_1 ... T_n$. The decision between theories is then intended to rest on n agreed-upon criteria $c_1 ... c_n$ that a theory ought (or ought not) to satisfy. The interlocutors agree to assign values on the closed real number interval $[-10, 10]$ representing the (positive or negative) importance of each criteria $w_{c_1} ... w_{c_n}$, and the extent $\mu_{c_j}(T_i)$ to which each theory T_i satisfies the criteria c_j. Then in the spirit of

Leibniz, they simply add the weighted extents together to provide the *rationality index* $\rho(T_i) \in [-10n, 10n]$ of each theory T_i:

$$\rho(T_i) = w_{c_1}\mu_{c_1}(T_i) + \ldots + w_{c_n}\mu_{c_n}(T_i)$$

Priest does not require that people perform this calculation as such, but emphasises that this is at least an ideal representation of how the debate ought to have played out:

> Note that I am not suggesting that in real-life disputes people actually sit down and do the calculations. Rather, the point is that when rational disputes are in progress, the arguments deployed may be understood as implicitly addressing the model. The model, then, gives a "rational reconstruction" of what actually happens.

The calculation is performed at a particular point in time and the interlocutors have convictions in common that allow them to determine how well each of the theories match each of the desired criteria.[2] So assuming the value of the weights w_{c_i} remain fixed, the calculation might change at a later time when new information changes the perspective of the interlocutors of how well the theories satisfy the criteria. Moreover, a new theory might be introduced for consideration. The model is in this way fallible.[3]

Priest intends WAM to provide a basis for a rational justification of revising logic, in particular to adopt a paraconsistent logic as the one true logic. Here consistency may be just one criteria, and a paraconsistent logical theory may perform better on some other criteria than a classical logical theory, enough to have a higher "rationality index". But Priest also concedes that WAM is open to logical pluralists:

> Even pluralists may debate which is the correct logic for a particular domain, application, etc. The methodology then applies. The debate between logical monists and logical pluralists is, in fact, a meta-debate, and we evaluate the two theories involved in exactly the same way.

Priest does not leave pluralist logical theories out of WAM, he simply interprets them as coming out worse:

[2] Presumably the set of convictions may be a set of propositions that the interlocutors know or believe, but Priest does not spell this out.

[3] Priest, at least within the bounds of the paper, does not appear to have anything to say about when the criteria themselves should change, though he does concede that the criteria and their weights are contentious issues.

> Unity is itself a desideratum; conversely, fragmentation is a black mark. Just think how one would react to an account of planetary dynamics which mooted quite different theories for each planet.

WAM in and of itself does not explain very much about how Priest regards paraconsistent logic as prevailing over classical logic, or how Priest rejects logical pluralism. The setting of criteria and the weights provide virtually all the substance of the argument.[4] Interlocutors who differ on the criteria might come to the opposite conclusions as Priest, while following WAM perfectly. What WAM does provide is, as with Williamson, an anti-exceptionalist view of logic, that logic is simply a theory like any other and does not have a privileged position with respect to the sciences.

Indeed, Timothy Williamson's abductive methodology, leading him to favour classical logic, can be compared with Priest's criteria for WAM. What constitutes a good abductive argument is a contentious matter, but thankfully Williamson provides an account of his own views in [91]. Theories, and in particular logical theories, are judged according to their:

1. Strength (Considered by Williamson to be the theory's informativeness)
2. Simplicity
3. Elegance (Is the logic axiomatisable? Is it finitely axiomatisable?)
4. Not entailing what can be falsified
5. Entailing a significant portion of what we can verify independently

Let us see what bearing this has on a proposed foundation of mathematics, between ZFC and some proposed paraconsistent set theory, say that of [87], which allows for a naive comprehension scheme at the expense of the expressivity of the conditional operator. Allowing for the naive comprehension scheme may be intuitive in some sense, but does not make paraconsistent set theory more simple or elegant as Williamson would have it. But it falls behind ZFC in strength and what it can entail, because modus ponens does not hold in general. So assuming that strength, simplicity, elegance, and entailment are given positive weight, ZFC is bound to prevail over paraconsistent set theory in the abductive methodology.

As one might expect, Priest's criteria are bound to be different from Williamson's, though he stops short of providing a definitive list. Some that he suggests include:

1. Adequacy to the data

[4] A detailed summary of Priest's arguments for adopting a paraconsistent logic can be found in [55].

2. Simplicity
3. Consistency
4. Unifying power
5. Avoidance of ad hoc elements

Let us return to our comparison of ZFC and paraconsistent set theory. For Priest, there is a strong prima facie case that set theory is not consistent, as naive comprehension is a close fit with our intuitions, which is what for him is meant by adequacy to the data. To modify naive comprehension in order to have consistency is thus a departure from our intuitions and an ad hoc move. What is gained from one criterion (consistency) is lost in two criteria (adequacy and avoidance of ad hoc elements), and it is clear that for Priest the criterion of consistency would not be weighted enough to compensate.

It is here that the differences between the two methodologies become apparent. Priest views the formulation of a logical theory as a modelling of an ideal rational intuition of validity, and naive comprehension and the T-schema (and their inconsistency) are for him rational and intuitive. Williamson treats logic as being free of metalinguistic content, and values logic in terms of its strength.

Despite appearances of a common ground in anti-exceptionalism Priest and Williamson are practicing two different disciplines, following different epistemic norms to reach different conclusions. There may be a criteria of deciding between different methodologies in WAM in terms of their relevance to particular epistemic norms. It is on these grounds that logical pluralism can be maintained.

2.4 Metatheoretic Perspectives on Abductive Reasoning

Ole Thomassen Hjortland's basis for a logical pluralism in [35] is also founded on an abductive methodology, but both he and Priest aim to obtain an ideal rational notion of validity. The abductive methodology is, unlike Williamson's, *metatheoretic* in its application. But then their methodology is also not *logic-neutral* in the sense that there is always some metalogic that informs the abductive methodology, which doesn't sit well with the tenets of anti-exceptionalism. In this section, this point will be elaborated.

The most important objection Hjortland has to Timothy Williamson's abductive methodology argument for classical logic is to the implicit embrace of a deflationary view of logic. And within the deflationary view is the assumption that statements of logical theories are universal generalisations. For Hjortland, logical rules have a "non-deflationary formulation": statements about validity expressed within a metalogic. Within this metalogic, one can speak of statements that are valid in a manner similar to relativisation within a class in set theory. For instance, Hjortland speaks of double negation elimination holding in PA, but not holding in the extension PAT of PA with a transparent truth predicate, in the following way:

$$\forall x : (Sent^{PA}(x) \to Val(\dot{\neg}\dot{\neg}x, x))$$

$$\neg\forall x : (Sent^{PAT}(x) \to Val(\dot{\neg}\dot{\neg}x, x))$$

Here *Sent* is the standard predicate on natural numbers x that holds whenever x is the code of a sentence. Accordingly $Sent^{PA}$ holds whenever x is the code of a sentence of PA, and $Sent^{PAT}$ holds whenever x is the code of a sentence of PAT. $Val(x, y)$, the property of validity, holds whenever x and y are codes of sentences ϕ and ψ such that given ϕ there is a valid proof of ψ.[5]

Hjortland views the unrestricted truth predicate and classical reasoning as prima facie desirable, but the two are incompatible. He takes this to be a blow against unrestricted generalisations and a case for logical pluralism. This move might be too fast. A pressing issue with Hjortland's approach is that the "deflationary" nature of the logical theory depends on the point of view of the reasoner: his approach is only "non-deflationary" with respect to PA and PAT, not with respect to the metalogic where universal quantifiers are indeed used. What is the justification for the choice of the metalogic? If we respond by accounting for the validity of the metalogic, we enter a regress.

For their part, Hjortland and Priest both acknowledge that their abductive methodology is not logic neutral. As Priest states, "In a choice situation, we already have a logic/arithmetic, and we use it to determine the best theory–even when the theory under choice is logic (or arithmetic) itself". Clearly they see this as something an anti-exceptionalist can live with, but it is not, as the regress argument shows that *some* metalogic has not been justified by the methodology. Perhaps it was chosen arbitrarily, which would leave us guessing if we made the right choice. Indeed, the metalogic might have been chosen on the basis of instrumental desires, ultimately prior to experience, contrary to the tenets of

[5] Hjortland speaks of the "Gödel code" of a sentence without further elaboration. The coding of a possibly uncountable number of sentences for any one of a possible uncountable number of logical languages cannot be done in the natural numbers. But with an ordinal numbering for each of the logical languages under consideration, ordinal-valued codes can be distinguished with a suitable pairing function.

anti-exceptionalism.

This issue at first may not appear to be a problem for Priest, who in [56] (for example) argues that the "same logic must be used in both 'object theory' and 'metatheory'". But Priest is compelled to adopt a very general notion of what sort of evidence a logical theory (and thus a logical metatheory) needs for a "best fit".

Priest refers to the evidence that a logical theory must account for as "our [informed] intuitions about the validity or otherwise of vernacular inferences . . . [and] forms of inference". There are, however, vernacular inferences involved in ensuring that the theory fits the criterion. How is the validity of those inferences to be assessed? Presumably by Priest's abductive criterion, but then we enter another regress.

The issues Priest and Hjortland have with self-applying their own methodology have the same cause: they both employ a priori assumptions about the nature of logic, which come prior to the implementation of the abductive methodology or (in Priest's case) are embedded in it, thus they are not anti-exceptionalists about logic. Moreover, what sets them against the deflationary view of logic already keeps them from being anti-exceptionalists.

2.5 Intuitionism and the Norm of Strength

For Timothy Williamson, the use of abductive criteria has led him to a unique "best", a form of classical logic, absent further evidence that would sway him to the contrary. If one followed Williamson's methodology as far as it goes, one might well be swayed to adopt classical logic as well, but as we will see in this section there is reason to doubt some of the criteria he adopts. One particularly powerful criterion that will be the focus of this section is the *norm of strength*: given two logical theories of otherwise about equal weight, prioritise the one of greater logical strength, i.e. the one that proves more theorems.

Objections to the norm of strength already exist. Stephen Read has already objected to this norm in [61] on the basis that Abelian logic is (like classical logic) post-complete, but provides accounts of conditionals that are, on the face of it, false. Yet there's a way out for Williamson here, as fit with the evidence was indeed one of his abductive criteria. Gillian Russell in [65] provides a notion of *scientific strength* of logics in terms of the expressivity of "subvalid" forms of argument as a means of refuting Williamson's more coarse-grained use of logical strength.

There is another objection, of greater relevance to the defence of the previous chapter's logical pluralism: the invocation of the norm of strength leads us to throw away intuitionistic logic, despite the epistemic benefits of having intuitionistic mathematicians around. The upshot of Williamson's argument is that we should all be classical mathematicians, to the likely detriment of progress in (at least) theoretical computer science and proof verification.

This objection is completely at odds with a defence of classical logic that Williamson has already written (see [94]): his challenge to deviant logics is the widespread applicability of classical mathematics, for classical mathematics "constitutes by far the most sustained and successful deductive enterprise in human history". This defence rests on a wholesale rejection of logical pluralism, which he has maintained in [92] and was already responded to in the previous chapter. But in any case, as we shall see, Williamson's defence of classical logic as being widely applicable fails when the role of logic in rational practice is adequately conceived.

The strategy Williamson uses is to put the burden of proof on the deviant logician. In order to revise logic, one must provide evidence for the usefulness of the revised logic in *every* situation, and show that the revised logic would be a greater benefit to all scientific inquiry than the previous. And, he argues, in order to do this one must rebuild mathematics from the ground up in the revised logic and make use of all of its useful applications, which for radical enough shifts (for instance, shifting from classical to paraconsistent logic) would be a very difficult endeavour. The stakes are too high to demand anything less. Legal cases hinge on reliable indications of mathematical certainty, bridges must be assured of their holding together, and so on.

In order for Williamson's argument to work, one must appeal to the universality of mathematics: the domain of quantification must be every single object of science. The quantifiers may be relativised to a particular appropriate subset for a given situation, but a change in the logic would change the way we interpret any scientific question and solve any problem.

Recall the previous chapter's motto on the role of logic: logic is a formalisation of a rational practice. But the practice might itself serve a different inferential role: intuitionistic mathematicians could be thought of as constructing mathematical structures from some basic elements, while classical mathematicians are trying to find them in a Platonistic universe. What to make of building bridges and projecting rocket flight paths in the respective paradigms is not out of the question in either case; the question is not what subset (of the domain of all of science) one is in, but what kind of activity one is willing to do.

The question Williamson might ask is how the activities are interpreted: perhaps as differing collections of axioms and rules of inferences, with the same first-order predicate language, ultimately formulated in a *classical metatheory* which try as he might the deviant logician cannot escape. This is what the classical logician does if he wants to see to what extent he can talk coherently about non-classical inferences inside his own language, but this is not how logicians of different traditions do or should understand one another; both parties would end up working in cross purposes. Instead, as mentioned in the discussion of Quine, it would make more sense of them to make use of interpretive charity with respect to each-other's practice.

Intuitionistic logic owes itself particularly well to interpretive charity because it has already been shown to extend to its own domain of mathematics with its own distinct kind of applications. And while this is just one example of a branch of mathematics whose underlying logic formalises epistemic norms that certainly contravene Williamson's norm of strength, this may be the *best* example, having been more widely applied than (say) any nonclassical theory of truth. There are indeed many conceptions of intuitionistic mathematics, but as Arend Heyting in [34] puts it, they tend to agree on many of the most important points, certainly enough to serve the present discussion.

Two of these points can be reformulated in terms of epistemic norms, which when put together contravene the norm of strength but can be shown to have their own benefits.

The first norm is the *proof-theoretic norm of meaning*: if you want to know the meaning of a proposition A, you must have a means to prove or refute A, and if A is a compound formula then its proof (refutation) is explained in terms of the proof (refutation) of the relevant constituent subformulas.[6] The second norm is the *constructive norm of meaning* of mathematical assertions: if you want to know the meaning of a mathematical assertion, you must know the means of construction of all the mathematical entities that the assertion claims exist.

The proof-theoretic and constructive norms of meaning are in general incompatible with the norm of strength as applied to logical theories, as intuitionistic proofs exclude the use of the law of excluded middle. The advantage of adopting them is, in brief, that one can provide witnesses to any assertion in mathematical practice. Taking an example from measure theory formulated in this way, the assertion that a set is of full measure implies that there is a procedure for finding a

[6]Michael Dummett and Dag Prawitz developed a different foundation for intuitionistic logic on the basis of a theory of meaning, fleshed out in the "justificationist semantics" seen in [21], but this is in a similar spirit.

point in that set (for a treatment see [9]). Intuitionistic proofs are also amenable to verification, as seen from work in [18] on the Calculus of Constructions.

It was not necessary for the purposes of this essay to argue that accepting proof-theoretic and constructive norms of meaning is necessarily preferable to accepting the norm of logical strength. But it is by now clear that accepting the proof-theoretic and constructive norms may be part of an epistemically productive mode of mathematical practice, providing insights that don't come from accepting the norm of logical strength.

There is a counterargument from *maximal productivity*: the vast majority of mathematics has been worked out through classical methods. However, the merits of classical mathematics is not what was under question, but rather the merits of Williamson's abductive criteria for revising logics. And one who has already accepted the proof-theoretic and constructive norms of meaning might have room to revise their choice of (intuitionistic) logic, but it won't include the norm of strength. The upshot of the preceding discussion is that there is not one epistemically productive mode of mathematical practice.

2.6 An Alternative Project of Logical Revision

In what follows, it shall be seen that on the conception of logic in the previous chapter, the use of general abductive criteria to revise logic is illegitimate, and the anti-exceptionalist project is not possible. This requires some reflection on what is involved in the revision of logical theories: that this revision occurs with the sole end of *narrow reflective equilibrium* between a practitioner's logic and the norms that constitute their practice, through deliberation on what propositions the practitioner is *committed* and *entitled* to by attribution in Robert Brandom's *game of giving and asking for reasons*. Such an approach is narrow enough in its aims and methods to preclude the use of abductive methodology such as Williamson's and Priest's.

Recall that the previous chapter's project of logical pluralism has the following approach:

- Through their genetics and past experiences, a practitioner has a set of *prior interests* that inform their choice of rational practice and *instrumental desires* that inform how they get at the truth. This is in keeping with Hilary Kornblith's contention in [39] that epistemic norms are informed by the "desires in a cognitive system that is effective at getting at the truth".
- A practitioner's choice of rational practice, together with their desires, lead them to a selection of *epistemic norms*, hypothetical imperatives contingent

on how well they fit their instrumental desires and prior interests, that constitute a guide for good practice. Some of the norms have more weight than the others in that they may be less amenable to revision.

- That practitioner *formalises* their epistemic norms into a logic.
- The logic entails consequences that (might) contravene some of the epistemic norms the practitioner already committed to.
- Depending on the weights the practitioner has assigned the norms, the practitioner (might) revise their epistemic norms with respect to their rational practice.

The context in which we can talk of interpretation of the logics of others is that there is a metalogical setting in which logics are interpreted and evaluated. This is best thought of as a social setting in which different practitioners interpret one another's assertions. But in revising logic and norms they do not merely interpret assertions, but also assess them.

The talk of desires being "in a cognitive system that is effective at getting at the truth" already gave us a story of how the norms and practices originated. But when evaluation is done on the basis of assessing whether a sentence is true, this poses a problem: on the one hand truths are supposed to be stable in some way, on the other logic should in principle be open to revision. Consider a deductive system K, which corresponds to a class of K-models. A familiar Tarskian picture would have us consider the set of sentences in the theory of every K-model to contain the "K truths". But we want to say there are circumstances in which we should revise K, the outcome of which may change that set of "K truths", suggesting that the "K truths" are not the set of sentences concerning K that are true simpliciter.

An elegant solution to the problem is to not talk of assessment in terms of whether another practitioner's statements are true, but more generally in terms of whether they are entitled to the claims they have implicitly or explicitly committed to. Such an assessment of the practices of others can be framed in Robert Brandom's *game of giving of asking for reasons* (see chapter three of [10], chapter six of [11]), a general framework of the speech acts of making and challenging assertions and how attributions of *commitments* and *entitlements* to their contents are affected by such acts.

The practitioner is assumed to be inclined to assert the epistemic norms that constitute their practice, which implies they are committed to them. (Thus, the scorekeeper's list of the other's commitments is always at least as large as the list of the other's assertions) The practitioner is also committed to the formalisation of the norms that they have chosen, and entailment is commitment-preserving.

The scorekeeper may assess that the practitioner is entitled to hold some of their commitments. But possibly not all of them, especially if they turn out to viciously conflict with one another. In Brandom's terminology two propositions are said to be *incompatible* if commitment to one precludes entitlement of the other.

Here Brandom's incompatibility semantics suggests an easy way of reformulating what it is we could mean by the truth while remaining agnostic about whether we should be realists about whether there is a single domain of truths or some ideal conditions under which we can gain access to them. Different practitioners are committed to the epistemic norms that constitute their practices, but they are also committed to the consequences of their logic, as the formalisation of those norms. They are also being evaluated by an interlocutor on whether they are entitled to their commitments. Truthful commitments are exactly those meant to withstand such scrutiny, but whether they would stabilise at some ideal limit of inquiry (or what that would look like) is outside the scope of this subject.[7]

Furthermore, recall the directive from the previous section:

> If you recognise a set of sentences Σ such that for each $Q \in \Sigma$, $P_1, P_2...P_n \vDash_K Q$, and you grant $P_1, P_2...P_n$, then you ought to see it to it that either every Q is granted, or that some of the norms invoked in granting $P_1, P_2...P_n$ or defining $\vDash_K$ are revised.

If a practitioner is committed to $P_1, P_2...P_n$, and they entail Q, the practitioner is committed to Q, but at times Q may preclude entitlement to another sentence of the language that the practitioner values, or preclude entitlement to an epistemic norm that underlies their practice. It is in these sentences that a scorekeeper may compel a practitioner to not necessarily grant every Q but to revise their norms or their formalisation. When such a scorekeeper can no longer challenge a practitioner's assertions on the basis of incompatibility of commitments to epistemic norms or their formalisation, then the practitioner's epistemic norms and their formalisation are said to be in narrow reflective equilibrium. (This is to be contrasted with *wide reflective equilibrium*, the subject of a later section)

This does not mean that logics cannot be criticised by any other means. For instance, one may question whether the practitioner is entitled to their stated

[7]Indeed, Brandom himself is strongly critical of any "ideal limit of inquiry" account, as in [11]: "there is no way to specify the ideality in question that is not either question-begging . . . or trivial". The outlined means of assessment does not contravene the existence of such an ideal limit of inquiry, but allows us to sidestep it along with direct talk of truth.

interests, desires, goals, or otherwise their practices in general. But these must be challenged directly (in terms of their incompatibility with some principle both parties are committed to, for instance), not from an account (alone) of the consequences of the practitioner's logic.

Logic depends on epistemic norms and their formalisation, which in turn depend on the goals, interests, and desires of the practitioner. This picture imposes a stratification of dependence, with interests, desires, and goals on a higher level than epistemic norms and their formalisation. The point is to demarcate the role of logic in theoretical revision as a whole in terms of this stratification: logic is to provide a framework for narrow reflective equilibrium on practices, as logical consequence tends to outrun our intuitions about coherence.

On this reading, the various problems with the anti-exceptionalist logical revision projects already discussed can be traced to an illicit use of abductive methodology as an evaluative device of logical and scientific theories together. Inference to the best explanation is there meant to be a means of evaluating epistemic norms and their consequences, but it does so on the basis of epistemic norms which are invoked in evaluating what a "best" explanation is. Logical and scientific theories aren't the same at all; scientific theories don't manifest themselves as commitments closed under a consequence relation, as they also come equipped with their own meta-level commitments of the goals meant to be pursued and the interests of the practitioner which themselves inform the meaning of the inference to the best explanation that Williamson might wish to employ.

So much for the doctrine of anti-exceptionalism about logic. In what remains, the importance of reflective equilibrium as a worthwhile procedure in epistemic evaluation shall be defended by appealing to the potential applications of each logic under consideration.

2.7 The Use of Logic for Reflective Equilibrium

In this section, the role for logic as establishing narrow reflective equilibrium is justified, against some principled objections.

Nelson Goodman introduced reflective equilibrium as an epistemic goal in [28] towards solving the problem of induction. Since then, Goodman's approach has been criticised by Stich & Nisbett in [76], Harman & Kulkani in [30], Paul Thagard in [79], and Jack Woods in [95]. The first point to make clear is that conditions for reaching reflective equilibrium are not to be regarded as providing sufficient conditions for adopting a logic, simply as one part of a larger picture, namely grounds for revising epistemic norms with respect to weights. The second point is that logical revision can in general lead us to an end-state of *narrow reflective equilibrium*; Jack Woods is perhaps right that asking for *wide reflective equilibrium* may be asking too much. In this light, some of the criticisms of Thagard remain salient, so they must be addressed.

Thagard characterises reflective equilibrium as being "at best incidental to the process of developing normative principles" and ultimately dispensable. While rejecting foundationalism in epistemology in principle, Thagard has over the years produced different approaches to a coherentist approach to the development of epistemic norms. For example, in [80] he advocates the following strategy:

1. "Identify a domain of practices."
2. "Identify candidate norms for these practices."
3. "Identify the appropriate goals of the practices in the given domain."
4. "Evaluate the extent to which different practices accomplish the relevant goals."
5. "Adopt as domain norms those practices that best accomplish the relevant goals."

This consequentialist approach to epistemic norm formation is entirely compatible with the present account: it fills in a number of details of how goals and instrumental desires establish epistemic norms and weights, such as in the identification of candidate practices and norms and in their matching with goals, whose soundness is not in question here. The interesting detail is that Thagard identifies his candidate processes as a means to replace reflective equilibrium entirely. This is because practice, appropriately chosen, is by his lights already coherent. But this puts too much faith in expertise: even skilled practitioners do not see the consequences of their own first principles.

The point of narrow reflective equilibrium as accounted for here is the establishment of *internal coherence* before a practice and its logic can be evaluated according to external criteria such as its usefulness. This is opposed to *wide reflective equilibrium* which takes into account external criteria: it is reached when a logic and the norms of the practice it formalises are compatible with themselves and with the norms of the community's relevant epistemic practices, particularly those that manifest in "background theories". Thagard in [79] argues against adopting narrow reflective equilibrium for establishing rational practice:

> Logical practice has improved enormously with the developments in deductive, inductive, and practical logic of the past several hundred

> years. In contrast, linguists do not aim to improve the overall grammar of a linguistic population, since their task is descriptive. The logician, on the other hand, is concerned to develop a set of principles that is inferentially optimal given the cognitive limitations of reasoners. This requires reference to background psychological and philosophical theories and to the goals of inferential behavior. Hence logical principles could only be arrived at by a process of wide reflective equilibrium.

Thagard thus appeals to the necessity of the logician to "develop a set of principles that is inferentially optimal given the cognitive limitations of reasoners", which requires the cultivation of background psychological theories. Yet being restrained by cognitive limitations in that respect is not necessarily a virtue of a logical theory, as we already saw in the previous section. By the standards already chosen here, by contrast, it is only demanded that the epistemic norms attain coherence with themselves. The reason for this limitation is that interpreting and at the same time evaluating the theories of others against our own would be an interpretive problem of a much greater magnitude, arguably an intractable one, as the interpretation of the other's practice would generally change with revision of one's own. Thagard asks too little of logic and too much of the role of reflective equilibrium.[8]

One other objection is anticipated. A practitioner chooses a logic as the best fit with their rational practice, but this leaves room for skepticism. The logic the practitioner chooses is, in practice, merely the best fit that they know. It may be that there are heretofore unknown logics that are themselves more faithful means of formalisation, "better fit" logics. Furthermore, it may be that some of these "better fit" logics would have lead the practitioner to choose different, contradictory, epistemic norms in their resulting state of equilibrium. Thus the state of equilibrium is contingent on our knowledge of formalisation.

The challenger would go on to say, "our particular state of equilibrium should not be contingent on our knowledge of formalisation, so logic should not be used to establish it." This is precisely where the present account diverges: formalisation provides us with the means to outrun our cognitive facilities through logical consequence, the alternative would be mere conjecture though which we would not

[8]Thagard shares Gilbert Harman's concerns in [29] about the normativity of logic: "The study of reasoning has always been a centerpiece of philosophy, but formal deductive logic captures so little of what is interesting about reasoning that it would be a grave mistake to take it as paradigmatic. At best, logic only tells you what you may infer from a given set of premises, not which of the infinite set of consequences you should infer." It is already clear where the present account differs from Harman on the question of what logic tells practitioners: to recap, the formalisation of epistemic norms into logic allows practitioners to outrun their own cognitive facilities of inference.

have certainty that equilibrium has been achieved. Even competent practitioners may fail to establish sound informal guidelines for their own practice.

2.8 Conclusion

Alfred Tarski in [77], an introductory textbook on logic, narrowed the role of logic to analysing “the meaning of the concepts common to all the sciences” and providing the laws that govern these concepts. This is in line with traditional monist accounts of logic as a tool for all scientific inquiry, concerns about the legitimacy of the enterprise of conceptual analysis notwithstanding. Yet at the same time, Tarski doubted whether there was a special “logic of the empirical sciences”, as the methodology of the empirical sciences had (by 1940) not progressed at the rate of the empirical sciences themselves. 79 years later, there is still no scientific consensus on an axiomatisation of all of physics, as Hilbert’s sixth program demanded. When Priest and Williamson speak of grouping logical theories together with scientific ones, governed by a single notion of rationality, they are already referring to a project that was always aspirational, even seeming more distant now than it was in the past. Experience has instead favoured a diversity of kinds of rational practice.

There is an alternative way to view logic, as the presentation of a guide to *some* practice that is constituted by epistemic norms, norms which may be accepted on the grounds that they satisfy our desires in a system that is effective at getting at the truth in the sense we heretofore outlined. The revision of logic may be done on the basis of narrow reflective equilibrium, so outlined, without an eye on broader “scientific” considerations.

Chapter 3

On Indefinitely Strengthened Liars

Already mentioned in the previous chapter is that a central motivation of logical revision for the anti-exceptionalists has been *semantic paradoxes*, i.e. paradoxes involving the intuitive use of a truth predicate, such as the *liar paradox* and *Curry's paradox*. Our findings provide grounds for a family of *pragmatic* solutions to the liar paradox along the lines of the practitioner simply identifying the norms she wishes to keep, and changing her practice or formalisation in accordance with either keeping a truth predicate in her logical language or not, so either way viciously paradoxical statements do not appear in the extension of a truth predicate in this language. For in general, semantic paradoxes are not a primary concern, but an inconvenience that one runs into in adding a truth predicate to one's language.

Where the truth predicate is retained, the exclusion of paradoxical sentences is often achieved by a regimentation of the extension of the truth predicate, typically using a construction like Saul Kripke's in [42].[1] There is an additional difficulty, the subject of this paper, on a common challenge in the liar paradox literature known as the *strengthened liar* (or the *revenge liar*) which is not a fault with the partial truth definitions in and of themselves but to their worth as "solutions" to the paradox. The potential presence of strengthened liar phenomena is typically used to conclude that such a regimentation of the extension of the truth predicate does not solve the paradoxes but is merely the product of ad hoc tricks, and a further regimentation in response to the strengthened liar phenomena leads to a vicious regress. This is in effect a challenge to the wide variety of pragmatic solutions that make use of this sort of regimentation to restrict the truth predicate's application, in favour of either a more all-encompassing attitude to logical revision (as presented by Graham Priest in [54]) or a revisionary attitude to the use of a truth predicate in general (as seen in Kevin Scharp's [67]).

[1] An alternative approach is using deductive systems such as Partial Kripke-Feferman in [36] to provide a corresponding regimentation of the rules that determine the truth predicate's use.

We will see what is meant by regimentation of this sort by refreshing our memory of Kripke's construction and the strengthened liar challenge to it. And in response, we shall see an example of how a seemingly vacuous strengthening of norms of truth may provide an underlying principle from which one can express strengthened liar paradoxes in nonclassical theories of truth on the basis of a family of *limit rules* for *revision sequences* of logical models. This will provide a thorough account of strengthened liars, while at the same time providing a case for the indefinite extensibility of truth value that Roy T. Cook has argued for. What's more, in this new framework "solutions" to the strengthened liar challenge can be reframed as "solutions" to the problem of how the concept of truth value (being *indefinitely extensible*) should be defined.

It is here claimed that by means of revision sequences of nonclassical models one can not only provide a faithful account of how a truth predicate functions in a logical language, but can also account for strengthened liar phenomena within the object language. In doing so, we can also account for what for the revision theory are limit rule discrepancies and strengthened liar paradoxes, to provide the beginnings of a more thorough alternative to Gupta and Belnap's approach where paradoxical sentences can be identified as such.

In the following sections we fix some definitions taken from the revision theory of truth in order to understand the *paraconsistent limit rule principle*, which will play a particular role: with this principle in mind, we find a different way to frame the strengthened liar problem. From the paraconsistent limit rule principle we will indeed find incidental support for the view of Roy T. Cook in [19] that the concept of truth value is indefinitely extensible. The ultimate upshot is that the "strengthened" or "revenge" semantic paradoxes, insofar as they can be formulated at all, do not provide a particular explanatory challenge to pragmatic solutions to the liar paradox that employ Kripke's construction.

3.1 Background

What is the liar paradox? Treatments of the *liar paradox* often begin with a formulation of the *liar sentence*:

This sentence is false. (L)

They typically follow this with an account of how this sentence is *paradoxical*, which in this context is to say that by seemingly innocuous reasoning it transpires that it is both true and false on the basis of its being either true or false. Say, suppose L were false. Then, since L refers to itself as being false, L is true. And suppose L were true. Then, since L refers to itself as being false, L is false. Thus reasoning with the liar sentence amounts to a paradox: the so-called liar paradox.

This is a fair place to start. But typically, such a beginning is followed by a "solution" for a problem that the liar sentence particularly poses for the formulation of truth in a logical language. In other words, much of the literature on the liar paradox has been motivated by the business of providing a definition of truth. We don't necessarily seek to find a definition of truth, but to show how to coherently express the liar sentence. This involves the imposition of a paraconsistent logic.

Much of the groundwork has already been done for us. In particular, the primary focus on this paper will be filling in an explanatory gap in a variation on Kripke's theory of truth, and seeing what it adds to our knowledge of the liar paradox. But first, it is worth noting how the perspective of thoroughgoing logical pluralism invites a pragmatic approach to semantic paradoxes, and an answer to the challenge that the "solution" to the liar paradox thus invoked is an ad hoc regimentation.

3.2 Kripkean Theories

In this section we provide a generalisation of Kripke's theory of truth from the manner in which it appeared in [42], defining what it means for a theory of truth to be *Kripkean* on the basis of its defining a truth predicate in terms of *jump operations*. We will start with a particular instance of Kripke's approach to the liar, then proceed to increase the level of generality to a notion of "Kripkean" theories of truth.

3.2.1 An Instance of Kripke's Theory of Truth

Kripke's theory of truth is a *semantic theory of truth*, that is, it provides an extension of a logical model (whose signature does not include a truth predicate) to a logical model with a truth predicate. But it is more accurate to say that Kripke's theory of truth is not a single semantic theory, but rather a family of semantic theories. Each theory belongs to a family of theories that differ in various ways. But we will start by providing a single well-motivated theory in this family.

The intuition that Kripke wished to capture for his "theory of truth" happens to map well into one particular set of assumptions. That intuition can be suggested in the following way. Let's say we want to express truth as a predicate acting on (codes of) sentences of a language of first-order arithmetic. What are the true sentences of our new theory, including the truth predicate? We take for granted that the *arithmetic truths* like $64 + 64 = 128$ are true, which in our case are just the sentences of our language of arithmetic that don't include the truth predicate, and the *arithmetic falsehoods* like $0 = 1$ are false. Knowing this, we in turn know that statements like "$64 + 64 = 128$ is true" are true, and statements like "$0 = 1$ is true" are false. Iterating the truth predicate and making use of logical combinations, we can also figure out that statements like "Either it is true that $64 + 64 = 128$ and ($2 \times 5 = 10$ is true), or $0 = 1$" are true, and statements like "It is both true that either $64 + 64 = 128$ or ($2 \times 5 = 10$ is true) and $0 = 1$" are false. In this way, we can start to build a partial definition of truth. Kripke proved that this iterative procedure, with a countable amount of iterations, leads to a *fixed point* where no more truths and falsehoods can be obtained.

However, we are carefully ignoring another class of sentences: those sentences which do not ultimately refer to arithmetic truths but instead to themselves, in a circular fashion. "This sentence is true" is one example, as indeed is the liar paradox "This sentence is false", but there are many others, such as "Either ($0 = 1$ is true) or this very sentence is true". We know those sentences exist because we know from the work of Alfred Tarski that first-order arithmetic with a truth predicate is sufficiently expressive to allow the formulation of self-referential sentences. But their truth value cannot be determined through the iterative procedure outlined above, so we simply omit these sentences from our definition of truth.

We now restate the demands to facilitate Kripke's theory. Let's say we want to add a truth predicate T to a model $\mathcal{M}'$ of first-order arithmetic (whose language $\mathcal{L}'$ has no truth predicate), with addition and multiplication, extending it to a model $\mathcal{M}$ whose language $\mathcal{L}$ contains T. And we want that truth predicate to be *transparent*, that is, for any sentence ϕ of the language $\mathcal{L}$ of $\mathcal{M}$, we want $\mathcal{M} \vDash T(\ulcorner\phi\urcorner)$ if and only if $\mathcal{M} \vDash \phi$, where $\ulcorner\phi\urcorner$ is the code of ϕ. The interpretation function for $\mathcal{M}'$ is classical, but $\mathcal{M}$ is a *paracomplete model*, that is, the interpretation function for $\mathcal{M}$ is three-valued with a third value $\varnothing$ for sentences that are considered neither true nor false. In particular we will employ the *Strong Kleene* or $K3$ semantics.[2]

We say that a formula ϕ is true in a Strong Kleene model $\mathcal{M}$ if with respect to any assignment to the variables it has semantic value $\{t\}$. And in accordance with Strong Kleene semantics, the conjunction/disjunction of a $\{t\}$-valued statement and a $\varnothing$-valued statement have truth value $\varnothing$ and $\{t\}$ respectively, the conjunction/disjunction of a $\{f\}$-valued statement and a $\varnothing$-valued statement have truth

[2]Alternatively, $\mathcal{M}$ may be seen as having a two-valued interpretation function that is only partial. We go ahead with the $K3$ semantics for the truth predicate as this will help facilitate a close comparison with the Logic of Paradox, which will follow in the next subsection.

value $\varnothing$ and $\{f\}$ respectively, the negation of a $\varnothing$-valued statement is $\varnothing$-valued, and conjunction and disjunction commute. We shall consider the Strong Kleene valuation scheme as given for the duration of this subsection.

Before we can define $\mathcal{M}$ and $\mathcal{L}$, we must define what we mean by the extension of a truth predicate.

3.2.1. Definition. A truth predicate T of $\mathcal{L}$ is said to have an *extension pair* $\langle T_t, T_f \rangle$, where T_t and T_f are sets of codes of sentences of $\mathcal{L}$. We say T_t is the *extension of the truth predicate* of $\mathcal{M}$, and we say T_f is the *anti-extension of the truth predicate* of $\mathcal{M}$.

Now we must define by transfinite induction an ordinal-valued sequence of partial models extending $\mathcal{M}'$, where the process of iteration that we gestured at earlier on takes real shape. But to do this we must first define, up to ω, an ordinal-valued sequence of corresponding extension pairs $\langle T_{t,\alpha}, T_{f,\alpha} \rangle$ to each model $\mathcal{M}_\alpha$ in the sequence:

- The *base model* $\mathcal{M}_0$ has a corresponding language $\mathcal{L}_0$ that extends $\mathcal{L}'$ with a truth predicate T with an extension pair referred to as the *base pair*. The base pair $\langle T_{t,0}, T_{f,0} \rangle$ of $\mathcal{M}_0$ is here assumed to be $\langle \varnothing, \varnothing \rangle$.
- Given an extension pair $\langle T_{t,\beta}, T_{f,\beta} \rangle$ of $\mathcal{M}_\beta$, the *jump operation* Φ on $\langle T_{t,\beta}, T_{f,\beta} \rangle$ is defined such that $\Phi\langle T_{t,\beta}, T_{f,\beta} \rangle = \langle T_{t,\beta+1}, T_{f,\beta+1} \rangle$ where $T_{t,\beta+1}$ is the set of all sentences assigned semantic value $\{t\}$ at $\mathcal{M}_\beta$, and $T_{f,\beta+1}$ is the set of all sentences assigned semantic value $\{f\}$ at $\mathcal{M}_\beta$ along with all well-formed formulas of $\mathcal{L}_\beta$ that are not sentences of $\mathcal{L}_\beta$. Then given $\mathcal{M}_\beta$, we define $\mathcal{M}_{\beta+1}$ to be the extension of $\mathcal{M}'$ whose truth predicate has the extension pair $\Phi\langle T_{t,\beta}, T_{f,\beta} \rangle$.
- Finally, $\mathcal{M}_\omega$ has extension pair $\langle \bigcup_{\alpha<\omega} T_{t,\alpha}, \bigcup_{\alpha<\omega} T_{f,\alpha} \rangle$, we call these instances *limit cases*.

It can be seen from the above that the extension of the truth predicate of $\mathcal{M}_1$ is exactly the set of arithmetic truths.

Kripke has proved that, for the valuation scheme that we have chosen, and the base pair of empty sets, the sequence of models will reach a fixed point; it is easy to see for our particular case that it is bounded above by ω. The key condition for a fixed point being reached is that the extension pair increases *monotonically*, that is to say that for all β, we have $T_{t,\beta} \subseteq T_{t,\beta+1}$ and $T_{f,\beta} \subseteq T_{f,\beta+1}$. The fixed point provides us with a definition of truth: given any sufficiently large α such that $\mathcal{M}_{\alpha+1} = \mathcal{M}_\alpha$, take $\mathcal{M} = \mathcal{M}_\alpha$ and define truth for first-order arithmetic to be the truth predicate of $\mathcal{M}$.

3.2.2 Some Variations of Kripke's Theory of Truth

We have just seen how, through iteratively defining truth over a partial model, we can define truth for first-order arithmetic without any paradoxes. But there are many other ways to do this. The ways can be divided into at least four kinds: the first by *changing the base pair*, the second by *changing the interpretation function* of the extensions of $\mathcal{M}'$, the third by *changing the limit cases* together with *the point at which truth is defined*, the fourth by adopting a *paraconsistent* approach.

The base pair need not be a pair of empty sets. Take ψ to be the sentence that says "this sentence is true", that is a sentence of $\mathcal{L}$ given by our diagonal lemma and defined such that $\psi : T(\ulcorner\psi\urcorner)$, where the colon indicates definitional equivalence. The base pair could be $\langle\varnothing, \{\psi\}\rangle$ or $\langle\{\psi\}, \varnothing\rangle$ with no difficulties: a fixed point model would still be reached, in either case. But if we take χ to be the liar sentence $\chi : \neg T(\ulcorner\chi\urcorner)$, if $\chi \in T_{n,0}$ for $n \in \{\{t\}, \{f\}\}$ a fixed point model will never be reached, and truth for first-order arithmetic will accordingly never be defined. If the base pair is a pair of empty sets, and a fixed point model is reached, the fixed point model is said to be the *minimal fixed point*.

The base pair variations are the least fundamental of the changes discussed here; they are said to define different fixed points with respect to the same interpretation function and limit and successor rules. Kripke defines metatheoretic notions based on differing fixed points. For instance, with respect to a paracomplete valuation scheme, a formula is said to be *grounded* if it is an element of either the set of true sentences or of the set of false sentences and non-formulas of the minimal fixed point. A formula is said to be *paradoxical* if it is an element of neither the set of true sentences nor of the set of false sentences and non-formulas for any fixed point whatsoever.

The interpretation function need not have the Strong Kleene valuation, there are others. Bas van Fraassen's supervaluations provide an alternative, as does the Weak Kleene valuation. Kripke's paper provides a thorough treatment that it would be too much of a digression to paraphrase here. It is also conceivable that the valuation scheme might range over more than three truth values.

Truth can also be defined over languages of infinite (but set-sized) cardinality, by defining the extension of the truth predicate by transfinite induction over all of the ordinals. Instead of having an ω stage, we have a similar definition for all limit ordinal cases.

The limit ordinal cases may be changed to accommodate a wider variety of valuation schemes, and truth might be defined as something other than a fixed point model. The textbook cases of this approach are in the revision theory of truth, worked out by Hans Herzberger in [32].

Finally, instead of working with paracomplete models, one might work in a "dual" fashion with paraconsistent models. Here is a brief example:

- Going back to our extension $\mathcal{M}$ of a model of first-order arithmetic $\mathcal{M}'$, have instead of a Strong Kleene interpretation function, an interpretation function that is valued according to the Logic of Paradox. Instead of having a third truth value $\varnothing$, the third truth value is $\{t, f\}$, and we say that a formula ϕ is true in a Logic of Paradox model $\mathcal{M}$ if with respect to any assignment to the variables it has semantic value $\{t\}$ or $\{t, f\}$, and false with respect to $\mathcal{M}$ if with respect to any assignment to the variables it has semantic value $\{f\}$ or $\{t, f\}$. The relations between the connectives are otherwise the same.

- In the transfinite induction sequence, the base pair may conventionally be defined as $\langle U, U\rangle$ where U is the set of all formulas of $\mathcal{L}$. As in contrast to the paracomplete models, where the idea is that we add truths and falsehoods to the extension pairs, with the paraconsistent models we shall take truths and falsehoods away.

- Given an extension pair $\langle T_{0,\beta}, T_{1,\beta}\rangle$ of $\mathcal{M}_\beta$, the *jump operation* Φ on $\langle T_{0,\beta}, T_{1,\beta}\rangle$ is defined such that $\Phi\langle T_{0,\beta}, T_{1,\beta}\rangle = \langle T_{0,\beta+1}, T_{1,\beta+1}\rangle$ where $T_{0,\beta+1}$ is the set of all sentences of $\mathcal{L}_{\beta+1}$ apart from those assigned semantic value $\{f\}$ at $\mathcal{M}_\beta$, and $T_{1,\beta+1}$ is the set of all formulas of $\mathcal{L}_{\beta+1}$ apart from those sentences assigned semantic value $\{t\}$ at $\mathcal{M}_\beta$ (including the formulas of $\mathcal{L}_{\beta+1}$ that are not sentences of $\mathcal{L}_{\beta+1}$).

- If γ is a limit ordinal, then $\mathcal{M}_\gamma$ has extension pair $\langle \bigcap_{\alpha<\gamma} T_{0,\alpha}, \bigcap_{\alpha<\gamma} T_{1,\alpha}\rangle$.

It can be shown that this construction reaches a fixed point model, which can be taken as a paraconsistent theory of truth extending first-order arithmetic. Indeed, given the choice of base pair, it is the *maximal fixed point* of the construction. Now with respect to a paraconsistent valuation schemes, a formula is said to be grounded if it is either not an element of the set of true sentences or not an element of the set of false sentences and non-formulas of the maximal fixed point. Also, a formula is said to be paradoxical if it is an element of both the set of true sentences and the set of false sentences and non-formulas for any fixed point whatsoever.

3.2.3 Some general definitions

There is a very general sense, outlined below, in which a theory of truth may be said to be Kripkean.

A jump operation is an operation on extensions of truth predicates on an ordinal-indexed sequence of models, called a *revision sequence*. The first model of a revision sequence, $\mathcal{M}_0$, is the *base model* of a language $\mathcal{L}_0$. Each model $\mathcal{M}_0$ includes at least one predicate T of a set $\mathcal{T}_0$ acting on codes of sentences of $\mathcal{L}_0$, whose extension is determined by the *extension tuple* $\langle T_{0,0}, T_{1,0}, ...\rangle$ of a sequence of *extension sets* of sentences of $\mathcal{L}_0$.

Given that some sentences of $\mathcal{M}_\alpha$ are assigned truth values $t_0, t_1 ...$, then the successor tuple $\langle T_{0,\alpha+1}, T_{1,\alpha+1}, ...\rangle$ is defined such that each $T_{\beta,\alpha+1}$ is the set of sentences of L_α that are assigned truth value t_β. Given a model $\mathcal{M}_\alpha$, the jump operation Φ on the tuple of extension sets $\langle T_{0,\alpha}, T_{1,\alpha}, ...\rangle$ is defined such that $\Phi\langle T_{0,\alpha}, T_{1,\alpha}, ...\rangle = \langle T_{0,\alpha+1}, T_{1,\alpha+1}, ...\rangle$. This is the *extension tuple* of $\mathcal{M}_{\alpha+1}$.

Any logical theory of truth that defines its desired notions of truth in terms of revision sequences, or fixed points of revision sequences, is here said to be Kripkean. This includes Kripke's theory of truth, Field's theory of truth, and all revision theories of truth that are defined in terms of models, as this notion of a "Kripkean" theory is sufficiently general to include them all. What is not included in this notion of a Kripkean theory is the definition of truth-at-a-model for M_α, the definition of each of the predicates in $\mathcal{T}_\alpha$ in terms of the extension tuple, and the definition of $\mathcal{M}_\lambda$ and $\mathcal{L}_\lambda$ when λ is a limit ordinal. This is where most Kripkean theories tend to differ.

The theories of truth we shall consider will from this point on all be Kripkean. We will continue with another illustrative example, Herzberger's revision theory of truth.

3.2.4 Herzberger's theory of truth

In this section we briefly outline the revision theory of truth, first outlined by Hans Herzberger in [32] and independently by Anil Gupta; we discuss the former, because it lends itself to a close comparison with Kripke's theory of truth with the Strong Kleene valuation scheme. We will then briefly discuss some criticisms of the approach.

Herzberger in [32] addressed a thought experiment: what would happen in Kripke's Theory of Truth if we were to start with a base model where all sentences are in the antiextension of T? Grounded true statements like "snow is white" (when expressible in the language) change from false to true after an iteration of Kripke jumps, and the truth teller sentences and grounded false statements stay false. Interestingly, for paradoxical sentences the truth values alternate periodically. What indeed does goes wrong is that the monotonicity of the jump operator fails to hold (due to the alternating truth values), so there are no fixed points, and there aren't even any sound models at the limit points in Kripke's inductive process. Herzberger addresses this problem with his own novel construction of models/languages, which he refers to as *semi-inductive* as opposed to inductive.

In Herzberger's semi-inductive construction, $\mathcal{L}_0$ may be defined such that T has empty extension, as before. After all, the antiextension of T in Herzberger's theory is precisely the complement of the extension, and the valuation scheme is also two-valued. Thus all sentences with a T-predicate are evaluated as false at $\mathcal{L}_0$. The successor case $\mathcal{L}_{\alpha+1}$ has the extension of T consisting of codes of true sentences of $\mathcal{L}_\alpha$; in the limit case $\mathcal{L}_\lambda$, $(T_{0,\lambda}, T_{1,\lambda}) = (\mathsf{liminf}_{\alpha<\lambda} T_{0,\alpha}, \mathsf{liminf}_{\alpha<\lambda} T_{1,\alpha})$ where for each limit ordinal λ $\mathsf{liminf}_{\alpha<\lambda}$ is defined as a partial function on ordinal-valued functions: $\mathsf{liminf}_{\alpha<\lambda} f(\alpha) = \{x : |\exists\delta < \lambda : \forall\gamma : \delta \leq \gamma < \lambda \to x \in f(\gamma)\}$.

Herzberger's approach comes with a proposed alternative to groundedness to identify paradoxical statements. That is, given a model $\mathcal{M}_\alpha$ in the semi-inductive hierarchy, a statement ϕ is said to be *stable at* $\mathcal{M}_\alpha$ in case either the code of ϕ is in $T_{0,\beta}$ for every $\beta > \alpha$, or the code of ϕ is in $T_{1,\beta}$ for every $\beta > \alpha$. Now ϕ is said to be *stable from* $\mathcal{M}_\alpha$ when it is stable at $\mathcal{M}_\beta$ for some $\beta > \alpha$. Also, ϕ is said to be *unstable from* $\mathcal{M}_\alpha$ when it is not stable at any $\mathcal{M}_\beta$ for which $\beta > \alpha$. A sentence is said to be *naively stable* in case it is stable from any choice of initial model $\mathcal{M}_0$ in the semi-inductive construction, and *naively unstable* in case it is unstable from any choice of $\mathcal{M}_0$.

The set of statements which are stable at $\mathcal{M}_\beta$, it is found, increase monotonically in magnitude: we can induce a function ψ on codes of these statements that is monotonically increasing, to replace ϕ. From this result, we find that Herzberger's theory of truth has an alternative to Kripke's fixed point theorem:

3.2.2. THEOREM. *From any model $\mathcal{M}_0$ at the start of Herzberger's semi-inductive construction, there is a model $\mathcal{M}'$ such that $\mathcal{M}' = \mathcal{M}_\lambda$ for some λ, and every statement that is stable from $\mathcal{M}'$ is stable at $\mathcal{M}'$.*[3]

Herzberger proves an important periodicity result of such constructions. In particular, there is an limit ordinal μ such that for each $T_{i,\beta}$ for $i \in \{\{t\}, \{f\}\}$, there is an ordinal $\alpha < \mu$ such that $T_{i,\beta} = T_{i,\alpha}$. Then after $T_{i,\mu}$, there is a *grand loop* of $T_{i,\gamma}$'s: for every γ such that $\gamma > \mu$, $T_{i,\gamma} = T_{i,\gamma+\rho}$ where ρ is said to be the *fundamental periodicity* of the semi-inductive construction.

Now $\mathcal{M}_\mu$ is a model such that every statement stable from $\mathcal{M}_\mu$ is stable at $\mathcal{M}_\mu$. But even better, one can determine the exact set of stable sentences: they are the sentences A for which either A is in the extension of T, or $\neg A$ is in the

[3]It is worth remarking that unlike in Kripke's theory, it does not matter what choice of the extension of T we pick at $\mathcal{M}_0$.

extension of T. Thus, $\mathcal{M}_\mu$ is a sound model with a two-valued truth predicate, for which a restricted version of Tarski's schema will be valid. We can consider it our desired M of the semi-inductive construction.

3.2.3. THEOREM. *From any model $\mathcal{M}_0$ at the start of Herzberger's semi-inductive construction, there is a model $\mathcal{M}_\mu$ such that μ is a limit ordinal, every statement that is stable from $\mathcal{M}_\mu$ is stable at $\mathcal{M}_\mu$, and a statement is naively stable if and only if either it is contained in $T_{0,\mu}$ or its negation is.*

So a theory of truth can be developed for what essentially is an adaptation of Kripke's construction, made to fit the classical valuation scheme.

The revision theory of truth has been praised for providing a comprehensive explanation of what statements in a logical language count as paradoxical with respect to a transparent truth predicate (given a particular valuation scheme), but criticised for not providing a definitive explanation of what sentences in such a language are true. Indeed there have been several different widely-cited proposals for a revision theory, generally differing in either the extension of T for the base model or in the limit rule.

Some of the most specific criticisms of the limit rules, concerning what sentences are declared stably true or false or unstable or neither, are among the revision theorists themselves (see [8] for instance). Other writers such as Hartry Field in [25] instead are broadly critical of revision theories of truth because of these various controversial limit rules: in particular, any revision theories of truth that have standard models of arithmetic are not compositional (see [44]).

Another move for the revision theorist to make is to accept that unstable sentences do not correspond to the world in the same way as stably true or false sentences do, and given that truth values are merely proxies of this correspondence, they should receive different truth values ([19]). The same, indeed, applies for sentences whose interpretation varies depending on the choice of starting point. Albert Visser's *extended Strong Kleene* revision theory in [84] provides an idea of what this could ultimately look like: paradoxical sentences considered "both true and false", and non-paradoxical sentences considered "neither true nor false". Nonetheless, as we now assign a new truth value we encounter a problem similar to Kripke's: the strengthened liar problem.

3.3 The Paraconsistent Limit Rule Principle

Recall that revision sequences of interpretations of sentences are essentially defined by an arbitrary initial hypothesis for the truth or falsity of that sentence and the limit rule that the revision theorist chooses. For their part, Gupta and

Belnap propose assigning unstable sentences an arbitrary choice of the classical truth values $\{t\}$ and $\{f\}$ at limit stages of the sequence. For the models at each of these limit stages there must be counterexamples to intersubstitutivity and compositionality, as each of the models are two-valued. For our purposes we shall propose an alternative limit rule. First we shall elaborate on these purposes, then propose the rule.

Hartry Field in [25] defines *the principle of intersubstitutivity of truth* and at least provides groundwork for *the principle of compositionality of truth*. Both are given as follows:

3.3.1. Definition. T is said to have *intersubstitutivity*, or follow *the principle of intersubstitutivity of truth*, if the following two conditions hold: (1) for any formula ϕ containing a subformula $T(\ulcorner\chi\urcorner)$, ϕ is logically equivalent to the formula ψ which is syntactically identical to ϕ apart from every instance of $T(\ulcorner\chi\urcorner)$ replaced with χ; (2) for any formula ϕ containing a subformula χ, ϕ is logically equivalent to the formula ψ which is syntactically identical to ϕ apart from every instance of χ replaced with $T(\ulcorner\chi\urcorner)$.

3.3.2. Definition. T is said to have *compositionality*, or follow *the principle of compositionality of truth*, if for any formulas ϕ and ψ, unary logical operator $\#$, binary logical operator $\odot$, the following three conditions hold: (1) $T(\ulcorner\phi \odot \psi\urcorner)$ is logically equivalent to $T(\ulcorner\phi\urcorner) \odot T(\ulcorner\psi\urcorner)$; (2) $T(\ulcorner\#\phi\urcorner)$ is logically equivalent to $\#T(\ulcorner\phi\urcorner)$; (3) $T(\ulcorner\forall x : \phi(x)\urcorner)$ is logically equivalent to $T(\ulcorner\phi(a)\urcorner)$ for any name a in $\mathcal{L}$.

We can think of these accordingly as norms that form part of a basis for a means of reasoning about the conditions under which statements are true within a logical language at least as expressive as first-order Peano arithmetic. They are, on their face, minimal conditions. Yet neither of these hold for Gupta and Belnap's formulations of the revision theory of truth, due to their reliance on classical models.

The limit rule plays a crucial part of the descriptive account that the revision theory is meant to provide, but for Gupta and Belnap, there is not a great difference between most of the limit rules that they find reasonable; they still "yield remarkably similar theories of definitions". But when finally reasoning in $\mathcal{L}$ within a classical model after T has been defined by the revision theory, as in Herzberger's theory, we do not have intersubstitutivity for T. We don't have compositionality for T either; if we did, then it is known that the theory would not have any standard models. What we end up with is something akin to the supervaluationist approach of Kit Fine, applied to truth; thought of as an *approximation* of how philosophers reason about truth, as each carries some artifacts.[4]

[4]Indeed Hartry Field makes this comparison in [25].

These artifacts are consequences of the underlying assumption that our theory of truth is classical.

If the sacrifices of compositionality and intersubstitutivity were necessary, a few exceptional cases might be a reasonable sacrifice. Yet as we are about to see this is not necessary, for one can conceive of Gupta and Belnap's goal as how one reasons from classical assumptions to paradoxes, and then reasons with paradoxes in nonclassical ways.

In Herzberger's theory of truth (see [32]) there is discussion of a peculiar property of revision sequences with a *constant* limit rule: that at some limit ordinal stage, a *grand loop* would be reached, beyond which unstable sentences would all return to the truth values from which they started. Thus no matter how large our language is, provided it is a set, there is a limit ordinal stage by which the unstable sentences of that language can all be determined. With this ordinal stage in mind, the *paraconsistent limit rule principle* is as follows: unstable sentences should be established *at that limit ordinal stage* to have all the truth values they were evaluated to have at each prior stage. We shall argue for this principle.[5]

A *truth value* of a sentence ϕ of $\mathcal{L}$ is conceived, as in [19], as being merely a proxy for the relation of ϕ to the world. And for our purposes of showing how we can reason about truth conditions, these relations are exhaustively provided by the different resultant revision sequences that ϕ may have.

If we find that from some bivalent initial interpretation ϕ is stably true (false), and from none is it stably false (true), then ϕ is interpreted as being $\{t\}$ ($\{f\}$). If ϕ is unstable from all bivalent initial interpretations, it has a different relation to the world, for we cannot say if it is or isn't true without saying it is both. To convey what that relation is, in keeping with conventions on the sense of a term we convey the process of how we got there; as ϕ alternates between true and false, it is ostensibly $\{t, f\}$.[6]

What is our desired valuation scheme for sentences of this new language with $\{t, f\}$? Is $\{t, f\}$ to be read as "both true and false" or "neither true nor false"? The attitude we shall take here is that as long as models with the truth predicate

[5]Thanks to Johannes Stern for a stimulating discussion on this topic. I can only hope that he finds the following rationale satisfactory.

[6]The work of this paper is only a small first step to exhaustively conveying the relation of ϕ to the world, if such a thing were even possible. One distinction that this analysis does not capture is that between unstable sentences that change truth value with every jump, and those that change truth value less frequently, which can be achieved by having a chain of mutually paradoxical sentences. Another is the distinction of truth-tellers from "no-no paradoxes" that contradict one another. A more thorough approach is left open.

T exist that are conservative extensions of those classical models without the truth predicate, and our initial demands of compositionality and intersubstitutivity for T are met, it does not matter. But for the sake of readability we shall assume for the remainder that our revision sequences are comprised of models for Graham Priest's *Logic of Paradox* (LP), as discussed in [52].

When in the revision process do we say that the liar has truth value $\{t, f\}$? As the statement of the paraconsistent limit rule principle alludes to, we can do it at a particular limit stage; at other limit stages we can assign the value $\{f\}$ to all unstable sentences, but it doesn't really matter what value among $\{t\}$ and $\{f\}$ we initially provide to paradoxical sentences at limits as the results will be the same. From a theorem of Vann McGee: we know that given $\boldsymbol{\kappa}$ is the successor limit cardinal of the maximum of the cardinality of $\mathcal{L}$ and ω, any stably true sentence is true at stage $\boldsymbol{\kappa}$. At this stage of the revision sequence, and at multiples of this stage, we can assign any unstable sentence the revised value: of all the truth values that it had been reasoned to have at prior stages.

For a complete theory of truth, we are not yet finished; we have not yet specified, for instance, what to do about the truth-tellers, or what (three-valued) valuation scheme we should choose for our jump operation. But the paraconsistent limit rule principle has broader applications when additional *strengthened liar* predicates are introduced, as in the following section.

3.4 The Strengthened Liar and a New Language

The Paraconsistent Limit Rule Principle provides a "true-and-false" truth value to the liar sentence. But when taken to further extremes, this rule can address the problem of strengthened liars wherever they may be expressed in the language. Before we can discuss how this works, we should be specific about what the strengthened liar problem is; there is some controversy on what constitutes a "real" strengthened liar. We formulate the strengthened liar and show how it can be treated using the Paraconsistent Limit Rule Principle, then formulate a language in which the strengthened liar (and its variants) can mainly be expressed.

Given that the liar sentence "this sentence is not true" is a means of diagonalising away from the truth values "true" and "false" to some other, in this case "true-and-false", what we mean by the *strengthened liar* is a statement that diagonalises away from all three of the truth values so mentioned, in this case "this sentence is not true, nor is it "true-and-false"". The difficulties do not arise within our paraconsistent theory itself, but rather in our classical metatheory. From assuming that the strengthened liar has truth value $\{1\}$ we can reason that it has truth value $\{0\}$ and vice versa, but from assuming that it has truth value

$\{1, 0\}$ we then reason that it has truth value $\{0\}$.

This broadly fits J.C. Beall's "third recipe" for a revenge theorist, which is paraphrased as follows:

1. Find a semantic notion; claim that it is in our "real" language;
2. Argue that that notion is not expressible in the logical language, on pain of triviality;
3. Conclude that the logical language is explanatorily inadequate.

For a family of paracomplete and paraconsistent theories of truth, including those under discussion, all that is needed for the above disaster is the addition of a new predicate or operator to the language that comes across as no more problematic than the truth predicate itself. In the spirit of keeping the logical operations untouched, we will work with new predicates: the first being $T_2(\ulcorner\psi\urcorner)$ meaning "ψ is "true-and-false" ". Paracomplete theories of truth face analogous strengthened liars such as "this sentence is neither true nor "neither true nor false" ", so they don't escape similar issues.[7]

Generally nonclassical theories of truth do not have the operators or predicates necessary to express strengthened liars, but the plausible case that these operators or predicates express coherent notions in and of themselves is often used to attack such theories as superficial and ineffective. There are a few ways in which the nonclassical truth theorist may respond:

1. The operators or predicates necessary to express strengthened liars do not express conceivable notions. (Seen in many approaches, but particularly [25])
2. The notion of truth under investigation does not hinge on truth values. (A possible defense of axiomatic approaches, such as in [36])
3. The metatheory is inconsistent; we should work on, for instance, a paraconsistent foundation of mathematics in which to define such troublesome notions as truth. (This has been a defense of Graham Priest at least since [54])

[7]Another alternative is that instead of predicates one may express strengthened liars in terms of combining new negation or conditional operators with the truth predicate. Philippe Schlenker in [68] adds negations ($\neg^1$, $\neg^2$, $\neg^3$...) to the language to express strengthened liars in a paracomplete framework, such as $\phi : \neg^2 T(\ulcorner\phi\urcorner)$ for "this sentence is neither true nor neither true nor false", or more succinctly "this sentence is not^2 true". Roy T. Cook has an analogous treatment with conditional operations. But Schlenker and Cook each impose many new truth values in the metalanguage, as we do, to both express and treat the strengthened liar problem.

4. Assign a new truth value to each strengthened liar. ([53], [68], [19])

The first three approaches would each run afoul of the goal we set out to achieve. We have already accepted "true-and-false" as a conceivable notion in the metalanguage, and with the paraconsistent limit rule principle entailing that certain sentences are true-and-false, we want to be able to simulate how a philosopher may reason with such a notion. We have thought of truth entirely in terms of models, in investigating truth as at least working like the relation between sentences/propositions and the world as Tarski conceived it. Finally, if we accept the metatheory as inconsistent, we solve a different problem.

Not only will we take the fourth approach, but more compellingly, we have already been led to it by the paraconsistent limit rule principle. Assuming we can express "true-and-false" by a second truth predicate T_2 as mentioned before, our strengthened liar is $\phi_2 : \neg T(\ulcorner \phi_2 \urcorner) \wedge \neg T_2(\ulcorner \phi_2 \urcorner)$. Applying the successor and limit rules prior to stage $\boldsymbol{\kappa}$, we see that ϕ_2 is unstable, so at $\boldsymbol{\kappa}$ it is assigned all the truth values it had before, thus it is assigned truth value $\{t, f\}$. But it is reevaluated as $\{f\}$ at stage $\boldsymbol{\kappa} + 1$, as the sentence says of itself that it is not "true-and-false". But at stage $2.\boldsymbol{\kappa}$, it is once again assigned all the truth values it had before, thus it is assigned truth value $\{t, f, \{t, f\}\}$.

One can conceive of yet more truth predicates, and corresponding strengthened liars for them: T_3 as "true and false and "true-and-false"" with liar ϕ_3, T_4 as "true and false and "true and false and "true-and-false""" with liar ϕ_4, and so on. How many truth predicates, and truth values, do we need? In [53] there are $\aleph_0$-many truth values, and for simplicity's sake we could stop there, but with enough semantic resources one can conceive of a liar ϕ_ω "this sentence is not in the extension of any of the truth predicates T_α where $\alpha < \omega$". The demand arises for truth values and truth predicates matching the entire class of ordinals, but then our language would not form a set. This forces us into proving not a global fixed point theorem but rather a *local fixed point theorem* for the collection of revision sequences from a nice (called *semantically-closed*) set of sentences of a nonclassical model (similar to that of [96], albeit localised as in [68]), which will be proved in the next section.

There is a distinct advantage in expressing revenge liars with predicates rather than operators. With only minor modifications to the proof in [53] one can show that we can have an entailment relation for such a many-valued paraconsistent logic, call it say LP', that is equivalent to that of LP. Take $\mathcal{L}$ to be the language of LP'. Before we can define each of the operations of $\mathcal{L}$, we need to fix notation for the truth values. The truth values other than $\{t\}$ and $\{f\}$ represent different "degrees" of paradoxicality, reached incrementally through the paraconsistent limit rule principle. It will help us later on to speak of an ordering, whereby greater degrees of paradoxicality of a statement are represented by a greater degree of brackets that the truth value has.

Now let us write out the ordering to define over the truth values that we obtain from our revision sequences: define a set K with elements $t_{\alpha,\beta}$, with α being any ordinal, and β being 0 or 1:

- $t_{0,0} = \varnothing$;
- $t_{1,0} = \{t\}$;
- $t_{1,1} = \{f\}$;
- $t_{2,0} = \{t, f\}$;
- For $\alpha > 1$, $t_{s(\alpha),0} = \{t, f\} \bigcup \{t_{\beta,0} : \alpha \geq \beta \geq 2\}$;
- $t_{\lambda,0} = \vee\{t_\alpha\}_{\alpha<\lambda}$.

3.4.1. Definition. Define the relation $>^t$ over elements of K by:

- $t_{0,0}$ is the least element of $>^t$;
- for $\beta = 0$ or 1, $t_{1,\beta} >^t t_{0,0}$ and nothing else;
- for $\alpha > 1$, $\alpha > \delta$, $\beta = 0$ or 1, $t_{\alpha,0} >^t t_{\delta,\beta}$.

3.4.2. Definition. $\geq^t$ is the reflexive closure of $>^t$.

The logical operations are defined as follows:

3.4.3. Definition. For each ordinal $\alpha \geq 2$, and $\mathcal{L}$-formula ϕ,
$\mathsf{I}(\neg\phi) = t_{1,0}$ whenever $\mathsf{I}(\phi) = t_{1,1}$;
$\mathsf{I}(\neg\phi) = t_{1,1}$ whenever $\mathsf{I}(\phi) = t_{1,0}$;
$\mathsf{I}(\neg\phi) = \mathsf{I}(\phi)$ otherwise.

3.4.4. Definition. $\mathsf{I}(\phi \wedge \psi) = t_{1,0}$ whenever $\mathsf{I}(\phi) = t_{1,0}$ and $\mathsf{I}(\psi) = t_{1,0}$; otherwise we have $\mathsf{I}(\phi \wedge \psi) = \mathsf{Min}\{\mathsf{I}(\phi), \mathsf{I}(\psi)\}$.

3.4.5. Definition. $\mathsf{I}(\phi \vee \psi) = t_{1,0}$ whenever $\mathsf{I}(\phi) = t_{1,0}$ or $\mathsf{I}(\psi) = t_{1,0}$; otherwise we have $\mathsf{I}(\phi \vee \psi) = \mathsf{Max}\{\mathsf{I}(\phi), \mathsf{I}(\psi)\}$.

3.4.6. Definition. $\mathsf{I}(\forall x_i\phi) = \mathsf{t}_{1,0}$ whenever for each sentence s, $\mathsf{I}(\phi[\mathsf{s}/x_i]) = \mathsf{t}_{1,0}$; $\mathsf{I}(\forall x_i\phi) = \mathsf{t}_{1,1}$ whenever for some sentence s, $\mathsf{I}(\phi[s/x_i]) = \mathsf{t}_{1,1}$; otherwise $\mathsf{I}(\forall x_i\phi]) = \wedge\{\mathsf{I}(\phi[s/x_i]) : \mathsf{s}$ is a sentence of $\mathcal{L}\}$.

Entailment is defined as follows. Take for LP' $\Sigma \vDash \phi$ if and only if for all interpretations I, either $\exists \psi \in \Sigma : \mathsf{I}(\psi) \notin D$ or $\exists \phi \in \Sigma : \mathsf{I}(\phi) \in D$, where D contains all and only the truth values t' for which $t \in t'$.

Our aim is that for $\alpha \geq 1$, we can define the truth predicates T_α over codes of sentences of $\mathcal{L}$ by:

- $T_\alpha(\ulcorner \phi \urcorner)$ if and only if $\mathsf{I}(\phi) = t_{\alpha,0}$;
- $\mathsf{I}(T_\alpha(\ulcorner \phi \urcorner)) = t_{\beta,0}$ if and only if $\mathsf{I}(\phi) = t_{\beta,0}$, where $\beta > \alpha$;
- otherwise we have $\neg T_\gamma(\ulcorner \phi \urcorner))$.

Indeed, we will show in the following section that we can provide a fixed point for a truth predicate over any set of sentences which is *semantically closed.* Moreover, any set of sentences can be shown to be a subset of some semantically closed set of sentences. The lesson being: if we want to know whether a sentence or any one of a set of sentences is true, false, or paradoxical, we can find out by forming a semantic closure over all of the sentences in question, and then establishing the truth predicate's extension. This is a context-dependent way of establishing truth.

3.5 Fixed Point Theorem

In this section, a method is outlined of reaching local fixed points of revision sequences for set-sized fragments of a first-order language with ordinal-many truth predicates over sentences of that language: a "vanilla" truth predicate, and others which express differing degrees of truth-and-falsity within a paraconsistent logic (the entailment relation from [53]).[8] To do this, we promote truth values of unstable sentences at certain conveniently large limit ordinals, according to the paraconsistent limit rule principle. The local fixed points are in the form of *alignment points* in the sense of Aladdin Yaqub in [96]. Yaqub demonstrates the existence of alignment points for revision sequences over an entire set-based language, with constant limit points, but nonetheless our proof will be similar to his.

The advantage of spelling out the ordering of the truth values in the previous section is that it allows us to more carefully formulate the local fixed point theorem, in terms of the *rank*, really the maximum of all the truth values by this ordering, of a set of sentences.

3.5.1. DEFINITION. Given that S is a set of sentences of $\mathcal{L}$, the *rank* of S is said to be the set $\mathsf{rank}(S) = \mathsf{Max}(\{\alpha : T_\alpha \phi \text{ is a constituent of } \psi \in S\})$.

[8] The truth predicates are in place of negations in [68].

In order to speak of local fixed points, we have to speak of local interpretations in terms of *semantically closed* sets of sentences, which are defined in terms of *constituent* sentences.[9]

3.5.2. Definition. For $\mathcal{L}$-formulas ϕ, ψ, we say that ϕ is an immediate constituent of ψ whenever either (a) ϕ is a subformula of ψ; or (b) ψ is $[\forall x_i : \chi(x_i)]\phi'$ or $[\exists x_i : \chi(x_i)]\phi'$, where ϕ is $\phi'[s/\chi(x_i)]$ for $\chi(x_i)$ a $\mathcal{L}$-formula with one free variable, s a sentence of $\mathcal{L}$.

3.5.3. Definition. The relation "ψ is a constituent of ϕ" is said to be the transitive closure of the relation "ψ is an immediate constituent of ϕ".

3.5.4. Definition. For S a set of $\mathcal{L}$-formulae, we say that S is *semantically closed* in i over $\mathcal{L}$ whenever: (1) for each formula ϕ in S, all of its constituents are also contained in S, and (2) for any formula $\phi \in S$ such that $\phi : T_\alpha(\ulcorner\psi\urcorner)$, $\psi \in S$.

3.5.5. Definition. The *S-fragment of* $\mathcal{L}$ is said to be the language $\mathcal{L}^S$ with only the truth predicates that exist in S.

We shall assume in the remainder of this section that S is a semantically closed set of sentences of $\mathcal{L}$ over i. This will ensure that for any paradoxical sentence, we can ascertain whether it is paradoxical or not with the assurance that its status will not change with new information.

3.5.6. Definition. The limit ordinal $\boldsymbol{\kappa}$ is defined by $\boldsymbol{\kappa} = \mathsf{Max}(\mathsf{card}(\mathcal{L}^S), \omega)^+$.

3.5.7. Definition. For some $\mathcal{L}^S$-formula ϕ in a particular model of $\mathcal{L}^S$, we say that the $t_{\alpha,\beta}$-extension of I is the set of all codes of sentences ϕ such that $\mathsf{I}(T_1(\ulcorner\phi\urcorner)) = t_{\alpha,\beta}$. We denote this set by $\mathsf{I}^{t_{\alpha,\beta},S}$ for each $t_{\alpha,\beta}$.

3.5.8. Definition. For each $\mathcal{L}^S$-formula ϕ, define the generalised jump operator f on the interpretation I by $T_\alpha(\ulcorner\phi\urcorner) \in \mathsf{f}(\mathsf{I}^{t_{\alpha,\beta},S})$ if and only if $\mathsf{I}(\phi) \geq^t t_{\alpha,\beta}$.

3.5.9. Definition. We say that ϕ is stable at a limit ordinal λ if there is an $\alpha < \lambda$ such that for all β with $\alpha \leq \beta < \lambda$, $\mathsf{I}_\alpha(\phi) = \mathsf{I}_\beta(\phi)$.

3.5.10. Definition. I_α is the function taking sentences ϕ of $\mathcal{L}^S$ to their truth value given the valuation of the revision sequence at α:

- Where $\gamma = 0$, I_0 is any two-valued extended $(t_{1,0}, t_{1,1})$ Strong Kleene valuation of S;

[9] The term is borrowed from [68] but the term "semantically closed" itself is misleading; this refers to a syntactic property of formulas.

- Where $\gamma = \alpha + 1$, $\mathsf{I}_\gamma(\phi) = \mathsf{f}(\mathsf{I}_\gamma(\phi))$;
- Where $\gamma = \delta \cdot \boldsymbol{\kappa}$ for some ordinal δ, and ϕ is stable at γ, $\mathsf{I}_\gamma(\phi)$ = the stablised truth value;
- Where $\gamma = \delta \cdot \boldsymbol{\kappa}$ for some ordinal δ, and ϕ is not stable at γ, $\mathsf{I}_\gamma(\phi) = t_{s(\alpha),0}$ where α is the greatest ordinal such that for some $\beta < \gamma$, $\mathsf{I}_\beta(\phi) = t_{\alpha,0}$;
- If γ is some other limit ordinal, then $\mathsf{I}_\gamma(\phi) = \mathsf{liminf}_{\alpha<\gamma}\mathsf{I}_\alpha(\phi)$

Finally we set about proving the alignment point theorem (Theorem 2).

3.5.11. Definition. For some limit ordinal α and semantically closed set S of sentences of $\mathcal{L}^S$, the tuple of sentences of S that are stable at α, U^S_α, is the tuple $\langle U_\alpha^{t_{0,0},S}, ...U_\alpha^{t_{\beta,0},S}\rangle$ where $\mathsf{rank}(S) = \beta$, and for $0 \leq \gamma \leq \beta$, and for each ordinal $\gamma < \mathsf{rank}(S)$, $U_\alpha^{\gamma,S} = \{\phi \in S : \exists \delta < \alpha : \forall \zeta : \delta \leq \zeta < \alpha \rightarrow \mathsf{I}_\delta(\phi) = t_{\alpha,\beta}\}$ is the set of sentences that are stable with a truth value in the $t_{\alpha,\beta}$-extension of α.

3.5.12. Definition. Ordinals $\alpha_1 ... \alpha n$ are said to be $\boldsymbol{\kappa}$-bounded whenever there is an ordinal γ such that for all $m \in \{1...n\}$, $\gamma.\boldsymbol{\kappa} \leq \alpha_m \leq (\gamma+1).\boldsymbol{\kappa}$.

3.5.13. Definition. For $\boldsymbol{\kappa}$-bounded ordinals $\alpha_1...\alpha_n$, an ordinal β is said to share $\alpha_1...\alpha_n$'s kappa bound when for all $m \in \{1...n\}$, there exists an ordinal γ such that $\gamma.\boldsymbol{\kappa} \leq \alpha_m \leq (\gamma+1).\boldsymbol{\kappa}$, and $\gamma.\boldsymbol{\kappa} \leq \beta \leq (\gamma+1).\boldsymbol{\kappa}$

3.5.14. Definition. The alignment point of the revision stages is a limit ordinal α for which the tuple of all sentences in a semantically closed set S that are stable at α is the same as the tuple of all sentences in a semantically closed set S that are stable for all limit ordinals $\beta > \alpha$ in the revision sequence.

3.5.15. Lemma. *For $\boldsymbol{\kappa}$-bounded limit ordinals α and β, if $U^S_\alpha = U^S_\beta$, then $\mathsf{I}_\alpha = \mathsf{I}_\beta$.*

3.5.16. Lemma. *For $\boldsymbol{\kappa}$-bounded limit ordinals α and β, and $\gamma < \boldsymbol{\kappa}$, if $\mathsf{I}_\alpha = \mathsf{I}_\beta$, then $\mathsf{I}_{\alpha+\gamma} = \mathsf{I}_{\beta+\gamma}$.*

Proof: By transfinite induction on γ:

- If $\gamma = 0$ then $\mathsf{I}_{\alpha+\gamma} = \mathsf{I}_\alpha = \mathsf{I}_\beta = \mathsf{I}_{\beta+\gamma}$.
- If $\gamma = \delta + 1$ for some ordinal δ, with $\mathsf{I}_{\alpha+\delta} = \mathsf{I}_{\beta+\delta}$, then taking the Kripke jump we have $\mathsf{I}_{\alpha+\gamma} = \mathsf{I}_{\beta+\gamma}$.
- If γ is a limit ordinal, then $U^S_{\alpha+\gamma} = U^S_{\beta+\gamma}$. But then from lemma 1, $\mathsf{I}_{\alpha+\gamma} = \mathsf{I}_{\beta+\gamma}$.

3.5.17. Lemma. *For every $\boldsymbol{\kappa}$-bounded limit ordinal α and β, if $U^S_\alpha = U^S_\beta$, then for every ordinal $\delta < \boldsymbol{\kappa}$, we have $\mathsf{I}_{\alpha+\delta} = \mathsf{I}_{\beta+\delta}$. And for every limit ordinal $\gamma < \boldsymbol{\kappa}$, we have $U^S_{\alpha+\gamma} = U^S_{\beta+\gamma}$.*

3.5.18. Lemma. *If α is a $\boldsymbol{\kappa}$-bounded limit ordinal, β a limit ordinal, $\beta < \boldsymbol{\kappa}$, and $U^S_\alpha = U^S_{\alpha+\beta}$, then for every $n < \omega$, we have $U^S_\alpha = U^S_{\alpha+\beta.n}$. Moreover, $U^S_{\alpha+\beta.n} \subseteq U^S_\alpha$.*

Proof: By induction on n:

- If $n = 0$ then we are done.
- If $n = m + 1$ then $U^S_{\alpha+\beta} = U^S_{\alpha+\beta(m+1)} = U^S_{\alpha+\beta.m+\beta}$. Since $U^S_\alpha = U^S_{\alpha+\beta}$, we have $U^S_\alpha = U^S_{\alpha+\beta.m+\beta} = U^S_{\alpha+\beta(m+1)}$.

Moreover, if $\phi \in U^S_{\alpha+\beta.\omega}$ then there is some $m < \omega$ for which $\phi \in U^S_{\alpha+\beta(m+1)}$. But we've just showed that $U^S_\alpha = U^S_{\alpha+\beta(m+1)}$, so $\phi \in U^S_\alpha$.

3.5.19. Lemma. *If α is a $\boldsymbol{\kappa}$-bounded limit ordinal, β is a limit ordinal, $\beta < \boldsymbol{\kappa}$, and $U^S_\alpha = U^S_{\alpha+\beta} = U^S_{\alpha+\beta.\omega}$, then $U^S_{\alpha+\beta.\omega} \subseteq \mathsf{I}_{\alpha+\gamma}$.*

Proof: Say $\phi \notin \mathsf{I}^{t_{\zeta,\zeta'},S}_{\alpha+\delta}$ for some ordinals ζ, ζ', and some $\delta < \beta.\omega$. For some $\mu < \beta$, and some n, $\delta = \beta n + \mu$. By the preceding lemmas $U^{t_{\zeta,\zeta'},S}_\alpha = U^{t_{\zeta,\zeta'},S}_{\alpha+\beta n}$, and $\mathsf{I}^{t_{\zeta,\zeta'},S}_{\alpha+\mu} = \mathsf{I}^{t_{\zeta,\zeta'},S}_{\alpha+\beta n+\mu}$. Indeed, then, $\phi \notin \mathsf{I}^{t_{\zeta,\zeta'},S}_{\alpha+\mu}$. From lemma 5 $U^S_\alpha = U^S_{\alpha+\beta m}$ for any $m < \omega$, so from lemma 3, $\mathsf{I}^{t_{\zeta,\zeta'},S}_{\alpha+\mu} = \mathsf{I}^{t_{\zeta,\zeta'},S}_{\alpha+\beta m+\mu}$. Then indeed $\phi \notin U^{t_{\zeta,\zeta'},S}_{\alpha+\beta.\omega}$, so $\phi \notin U^{t_{\zeta,\zeta'},S}_\alpha$.

3.5.20. Lemma. *If α is a $\boldsymbol{\kappa}$-bounded limit ordinal, β is a limit ordinal, $\beta < \boldsymbol{\kappa}$, and $U^S_\alpha = U^S_{\alpha+\beta} = U^S_{\alpha+\beta.\omega}$, then the set of sentences of $U^S_\alpha = U^S_{\alpha+\beta.\omega.\gamma}$ for any ordinal $\gamma < \boldsymbol{\kappa}$.*

Proof: By transfinite induction on γ:

- If $\gamma = 0$, we are done.
- If $\gamma = \delta + 1$, by a previous lemma, $U^S_{\alpha+\beta.\omega.(\delta+1)} = U^S_{\alpha+\beta.\omega.\delta+\beta.\omega} = U^S_{\alpha+\beta.\omega} = U^S_\alpha$.
- If γ is a limit ordinal, then if $\phi \in U^S_{\alpha+\beta.\omega.\gamma}$ then there is some $\nu < \gamma$ such that $\phi \in U^S_{\alpha+\beta.\omega.\nu+\beta.\omega}$, that is $\phi \in U^S_{\alpha+\beta.\omega.(\nu+1)}$. Now since γ is a limit ordinal, $\nu + 1 < \gamma$, so $\phi \in U^S_\alpha$, so $U^S_{\alpha+\beta.\omega.\gamma} \subseteq U^S_\alpha$. Moreover since for every $\mu < \gamma$ we have $U^S_{\alpha+\beta.\omega.\mu} = U^S_\alpha$, by the preceding lemmas we have $U^S_\alpha \subseteq I_{\alpha+\beta.\omega.\mu+\pi}$ for any $\pi < \beta.\omega$, thus $U^S_\alpha \subseteq U^S_{\alpha+\beta.\omega.\gamma}$.

3.5.21. Lemma. *If α is a $\boldsymbol{\kappa}$-bounded limit ordinal, β is a limit ordinal, $\beta < \boldsymbol{\kappa}$, and $U^S_\alpha = U^S_{\alpha+\beta} = U^S_{\alpha+\beta.\omega}$, $U^S_\alpha \subseteq I_{\alpha+\gamma}$ for any ordinal $\gamma < \boldsymbol{\kappa}$.*

3.5.22. LEMMA. *If α is a $\boldsymbol{\kappa}$-bounded limit ordinal, β is a limit ordinal, $\beta < \boldsymbol{\kappa}$, and $U^S_\alpha = U^S_{\alpha+\beta} = U^S_{\alpha+\beta.\omega}$, then $U^S_\alpha = U^S_{\alpha+\beta.\gamma}$ for every ordinal $\gamma < \boldsymbol{\kappa}$.*

3.5.23. LEMMA. *For all $\boldsymbol{\kappa}$-bounded limit ordinals α, β, if $\alpha < \beta$, there is some limit ordinal $\gamma < \boldsymbol{\kappa}$ such that $U^S_\beta = U^S_{\beta+\gamma}$.*

Proof: Suppose otherwise. Then for each such α, there is such a β with $\beta > \alpha$ and $U^S_\beta \neq U^S_{\beta+\gamma}$ for all such γ. Define α' to be the least limit ordinal such that $\alpha' > \alpha$ and $U^S_{\alpha'} \neq U^S_{\alpha'+\gamma}$ for all such γ. Define $f(\mu)$ for each μ sharing α's $\boldsymbol{\kappa}$-bound as follows:

- $f(\zeta.\boldsymbol{\kappa}) = \omega_1$ where ζ is the greatest ordinal such that $\zeta.\boldsymbol{\kappa} < \alpha$;
- $f(\nu + 1) = f(\nu)'$, $f(\nu)'$ defined like α';
- $f(\lambda) = \bigcup_{\nu<\lambda} f(\nu)$, ν a limit ordinal.

We can show that f maps to limit ordinals by transfinite induction over each ordinal μ that shares α's $\boldsymbol{\kappa}$ bound:

- If $\mu = 0$, then $f(\mu) = \omega_1$ which is a limit ordinal;
- If $\mu = \nu + 1$, then $f(\mu)$ has already been stipulated to be a limit ordinal;
- If $\mu = \lambda$, it is clear from the induction hypothesis that for each of the predecessors $f(\nu)$ of $f(\lambda)$, $f(\nu)$ is a limit ordinal. Since $\boldsymbol{\kappa}$ is an infinite successor cardinal, it is regular. Therefore $\bigcup_{\nu<\lambda} f(\nu) = f(\lambda)$ is also a limit ordinal.

Consider the set $\{U^S_{f(\nu)} : \nu$ shares α's $\boldsymbol{\kappa}$ bound $\}$. Distinct ν correspond to distinct elements of this set; indeed, this set has $\boldsymbol{\kappa}$ distinct members, which is impossible.

3.5.24. LEMMA. *There is a $\boldsymbol{\kappa}$-bounded limit ordinal α and a limit ordinal $\beta < \boldsymbol{\kappa}$ such that $U^S_\alpha = U^S_{\alpha+\beta} = U^S_{\alpha+\beta.\omega}$.*

Proof: We have already determined that there is a limit ordinal γ sharing α's $\boldsymbol{\kappa}$ bound such that for every limit ordinal δ sharing α's $\boldsymbol{\kappa}$ bound, if $\delta > \gamma$, then for some limit ordinal $\pi < \boldsymbol{\kappa}$ we have $U^S_\delta \neq U^S_{\delta+\pi}$. Take γ_0 to be the least such ordinal. We know that for any ordinal $\delta > \gamma_0$, there is an ordinal δ' for which $U^S_\delta = U^S_{\delta+\delta'}$. Take δ'_0 to be the least such ordinal. Now define $g(\pi)$, for all $\pi < \boldsymbol{\kappa}$, by:

- $g(0) = \gamma_0$, where γ_0 is the least limit ordinal sharing α's $\boldsymbol{\kappa}$ bound such that for all limit ordinals β sharing α's $\boldsymbol{\kappa}$ bound that are greater than γ_0, $U^S_\beta = U^S_{\beta+\gamma}$ $\zeta.\boldsymbol{\kappa} < \alpha$;

- $g(\mu+1) = g(\mu) + (g(\mu))'.\omega$, with $(g(\mu))'$ defined like δ';
- $g(\lambda) = \bigcup_{\mu<\lambda} g(\mu)$, for λ a limit ordinal.

We can show that g maps to limit ordinals by transfinite induction over each ordinal π that shares α's $\boldsymbol{\kappa}$ bound:

- If $\pi = 0$, then $f(\mu) = \gamma_0$ which has already been stipulated to be a limit ordinal;
- If $\pi = \mu + 1$, then $f(\mu)$ has already been stipulated to be a limit ordinal;
- If $\pi = \lambda$ a limit ordinal, then from the induction hypothesis for each of the predecessors $f(\mu)$ of $f(\lambda)$, $g(\mu)$ is a limit ordinal. Since $|\nu|.\boldsymbol{\kappa}$ is an infinite successor cardinal for each ordinal $\nu > 0$, it is regular. Therefore $\bigcup_{\mu<\lambda} g(\mu) = g(\lambda)$ is also a limit ordinal.

So g maps to limit ordinals. We show by transfinite induction on ξ that $U^S_{g(\pi+\xi)} \subseteq U^S_{g(\pi)}$ for ξ sharing α's $\boldsymbol{\kappa}$-bound:

- If $\xi = 0$ the result is immediate;
- If $\xi = \mu + 1$, then $g(\pi+\xi) = g(\pi+\mu+1) = g(\pi+\mu) + (g(\pi+\mu))'.\omega$ and from a previous lemma we have $U^S_{g(\pi+\xi)} \subseteq U^S_{g(\pi)}$;
- If $\xi = \lambda$ a limit ordinal, with λ a limit ordinal, then if ϕ is stable at $g(\pi+\lambda)$ then ϕ is stable at $g(\pi+\mu+1)$. Since $\mu < \lambda$, we have $\mu + 1 < \lambda$ since λ is a limit ordinal. By the inductive hypothesis, $\phi \in U^S_\pi$.

So the set of $U^S_{g(\pi)}$'s is linearly ordered by $\subseteq$, and since it is smaller than $\boldsymbol{\kappa}$ there must be at least two identical $U^S_{g(\pi)}$'s.

3.5.25. Theorem. *If there are limit ordinals α (which is $\boldsymbol{\kappa}$-bounded) and $\beta < \boldsymbol{\kappa}$ such that $U^S_\alpha = U^S_{\alpha+\beta} = U^S_{\alpha+\beta+\omega}$, then for each ordinal $\gamma < \boldsymbol{\kappa}$, $\alpha+\beta.\gamma$ is an alignment point of the revision stages.*

Proof: We have already found that for each $\boldsymbol{\kappa}$-bounded ordinal α, $I_{\alpha+\gamma}$ contains U^S_α. So sentences that are stable at α stay stable through to the next $\boldsymbol{\kappa}$ bound, and beyond. Suppose that ϕ is a sentence that stays stable through to the next $\boldsymbol{\kappa}$ bound. Then ϕ is stable at $\alpha + \beta(\mu+1)$ for some ordinal μ. From a previous lemma, ϕ is stable at α, thus it is stable at $\alpha + \beta.\gamma$.

3.5.26. Theorem. *$\boldsymbol{\kappa}$ is an alignment point of the revision stages of $\mathcal{L}^S$.*

Proof: The α and β of theorem 1 are smaller than $\boldsymbol{\kappa}$, and $\boldsymbol{\kappa}$ is a limit ordinal, so $\boldsymbol{\kappa} = \alpha + \beta.\boldsymbol{\kappa}$ is a special case.

3.5.27. Theorem. *Every sentence of S that is stable at $\alpha.\boldsymbol{\kappa}$ is also stable at $\beta.\boldsymbol{\kappa}$ for any ordinal $\alpha < \beta$.*

3.5.28. Theorem. *For any semantically closed set S of $\mathcal{L}^S$-sentences, the sequence of subsets S_β of S that consist of the sentences stable at the least $\alpha.\boldsymbol{\kappa}$ such that $\beta < \alpha.\boldsymbol{\kappa}$ is monotonically increasing.*

Below is the local fixed point theorem.

3.5.29. Theorem. *For any semantically closed set S of $\mathcal{L}^S$-sentences, there is some γ such that for all $\delta \geq \gamma$, $S_\gamma = S_\delta = S$.*

Proof: If not then S has as many distinct elements as there are ordinals, which is impossible because S is a set.

3.6 Indefinite Extensibility

The appeal of paraconsistent logics in the treatment of the liar paradox is that the liar sentence, along with all strengthened liar sentences, are prima facie both true and false. Avoiding this conclusion requires us to reckon with additional assumptions that turn out not to be necessary. Yet accommodating the semantic paradoxes comes with some metaphysical consequences: it has already been suggested at the beginning of this paper (and by Roy Cook) that PLRP provides support for the idea that the concept of truth value is indefinitely extensible.

Here we explain what we mean by indefinite extensibility, and how PLRP provides support for this. This is also relevant ground to elaborate on a soft objection to our approach, which amounts to suggesting that accommodating strengthened liars doesn't have to be as complicated as we made it out: there only need be $\aleph_0$ many truth values, so $\mathcal{L}$ might form a set. We show grounds for skepticism to this idea.

It is possible that one might conceive of fewer revenge liars than this paper argues. In particular, there is a perspective from which one might conceive of at most $\aleph_0$-many revenge liars. This would come about through believing in certain limits of conceptualising truth, analogous to finitism in mathematics. Say you believe that for a concept to be conceivable it must be definable in a form of an English sentence that is of finite length and with no reference to other non-primitive concepts. One might for example define the revenge liars in a recursive, increasingly long manner like this:

- The liar sentence is the sentence that says of itself that it is false.

- The first revenge liar sentence is the sentence that says of itself that it is false, and it is also not whatever one would classify "the sentence that says of itself that it is false" to be.

- The second revenge liar sentence is the sentence that says of itself that it is false, and it is also not whatever one would classify "the sentence that says of itself that it is false" to be, and it is also not whatever one would classify "the sentence that says of itself that it is false, and it is also not whatever one would classify 'the sentence that says of itself that it is false' to be" to be.

- ...

For any finite n, the nth revenge liar sentence is definable in this way. However, the $\aleph_0$ revenge liar sentence is not, as the definition could no longer be of finite length. If someone regards this as a sufficient argument that the $\aleph_0$ revenge liar does not exist, then there are only $\aleph_0$-many revenge liars. In that case, it would not follow from the Paraconsistent Limit Rule Principle that truth with every revenge-motivated truth predicate cannot be defined globally over the entire language; as long as the fragment of the language not containing the truth predicates forms a set then truth can be defined globally, as the language would have at most $\aleph_0$ many truth predicates and can thus form a set. The class of truth values would be indefinitely extensible, but it would form a set of cardinality $\aleph_0$. In short, the problem of how the revenge liar paradoxes may be expressed would still be solved, but the solution would be much easier than what we have done here.

There is reason to not just take for granted that there are merely $\aleph_0$ many revenge liar predicates. In order to exclude the transfinite revenge liars in a non-ad hoc way, we should have some principled reason to exclude the transfinite from our theory of definitions. Being restrictive about the conceivability of concepts to the above extent would be one way to do it, but this would exclude from our class of conceivable concepts a considerable amount of the mathematical concepts that one might hope to define.

3.7 Ultimate Revenge

The local fixed point theorem provides a basis for including strengthened liar predicates, but it is missing an account of a challenge that has arisen in the liar paradox literature many times. This is a notion we will call *ultimate revenge*.

The *ultimate revenge* predicate cannot be defined: this is the predicate F on sentences for which for any sentence ϕ, $F(\ulcorner\phi\urcorner)$ if and only if ϕ is false only. There is no truth value in the hierarchy we have defined for which $F(\ulcorner\phi\urcorner)$ is stable. The revenge theorist can present the inability to define F as the necessary proof that we have not really dealt with strengthened liars.

Roy T. Cook already suggested a way to respond without sacrificing a classical metatheory. The truth theorist could stipulate that the ultimate revenge liar cannot be expressed. The idea would be that the ultimate liar would be extensionally the same as the "limit" of the sequence of ordinal-valued predicates T^F_α on sentences for which for any sentence ϕ, $T^F_\alpha(\ulcorner\phi\urcorner)$ if and only if ϕ does not have the value true or any of the other revenge truth values up to t_α. There is no "limit" of limit ordinals, so perhaps the lesson of the ultimate revenge liar is that what it says is "the truth value of this sentence will never be reached", which is true but cannot be conveyed in the metalanguage.

Like many "solutions" to the revenge liar of this form, this line of argument is too presumptuous. The revenge theorist of Beall's recipes doesn't have to examine all the ordinal-valued truth predicates to come up with the ultimate revenge predicate. It's an option that's been available to her from the beginning.

Thankfully, closer examination reveals the ultimate liar can be treated in a similar way to the strengthened liars mentioned earlier. Recall that a truth value merely serves as a proxy for the relation between a sentence and the world. There should be nothing in principle against assigning a new sort of proxy to the sentences whose relation to the world is at least partly characterised by an inability to find a truth value. After all, we have already found support for the idea that "truth value" is an indefinitely extensible concept. The ultimate revenge liar could be given its own classification, in a new category of "deviation values" orthogonal to the existing truth values.

But in being assigned its own classification a new class of revenge liar paradoxes could be conceived of, in a more general form encompassing this new classification along with all of the truth values, and the same lessons would be learned as before. One might say we have entered a regress. But another way of putting it is that we have managed to characterise the revenge liar phenomena. The lesson from the revenge liar is that with sufficient expressive resources one can always diagonalise out from the truth values available, and this is a property of truth values as much as transparency is taken to be a property of truth.

3.8 Conclusion

This paper solves the problem of how the revenge liar paradoxes may be expressed for (paraconsistent) Kripkean theories of truth. But there are two lessons that come with this:

1. Truth (with every revenge-motivated truth predicate) cannot be defined globally over the entire language, but rather locally for set-sized fragments of the language;

2. The class of different truth values is indefinitely extensible and may not form a set. Roy T. Cook has already proposed this, but through the Paraconsistent Limit Rule Principle we see how this comes about.

The first lesson is another challenge to any view of logic as being a foundation of mathematics and formal semantics that at once makes use of universal quantifiers. After all, either $\mathcal{L}$ fails to include all of the strengthened liars, or quantification must be restricted to set-sized domains.

Many of the results may transfer to different nonclassical logics that a practitioner may be inclined toward for some reason. For instance, for many of the alternatives to LP with a stronger or otherwise different conditional operator, the above results should also carry over. Each paraconsistent logic like LP has a "dual" paracomplete logic (K3 in the case of LP), and many of the results would carry over in a dual form, where a sentence's having more than one truth value would mean that it was neither true nor false (or variations thereof, such as "neither true nor false nor 'neither true nor false'").

Chapter 4

Reflections on Knowing-How

4.1 Background and Introduction

Logic is said to be a formalisation of one's rational practice, as constituted by the epistemic norms we adopt and their consequences. We can then say that through applying logic we can learn and be certain of consequences of some of our norm-following behaviours, for instance from understanding the properties of numbers and the implementation of formal number theory we can verify proofs such as Selberg's proof of the prime number theorem [3]. But there are formal theories of arithmetic, for instance Peano Arithmetic and Heyting Arithmetic, that are substantively different enough that they are supposed to formalise different practices. We would like to be able to say what it is that, in particular, a classical formal mathematician can learn from a nonclassical formal mathematician, and vice versa–call this the *Problem of Communicability.* As we have already seen that the practice-oriented principle of charity is necessary in an account of how a solution to the Problem of Communicability is possible, it is natural to address the Problem of Communicability on the basis of a characterisation of the *potential ability* or *strong knowledge-how* for the formal mathematician to make inferences.

It is imperative to first rule out one approach to this characterisation, the *Propositional Thesis*: that the knowledge-how of a logician may be reduced to their knowledge-that of the propositions expressed by their logic's entailed sentences, or their knowledge-that of the propositions expressed by the sets of sentences that follow from others that they wish to accept as given. Then a pragmatics for a conversation between two practitioners is outlined.

4.2 Knowledge-How and Knowledge-That

There is, of course, a rich literature on the extent, characterisation, and paradoxes of what could be said to be propositional knowledge, including a wide range of formal treatments. Compared with this, there are relatively few formal treatments of knowledge-how, at least not explicitly as a predicate of a formal language; Yanjing Wang in [86] and Wang and Tszyuen Lau in [43] are included among the few, the former a goal-oriented approach whose characterisation of an interlocutor's knowledge-how does not depend on the agent's information state, the latter an attempt to characterise knowing-how in terms of knowledge-that of an ability. In light of this, it might be tempting to characterise the knowledge-how of a logician simply in terms of knowledge-that by means of the Propositional Thesis. In the following, that approach will be shown to be misguided.

Logic, in my conception, is a formal presentation of a guide to undertaking a rational practice, which is constituted by epistemic norms and their consequences. So a rational practice, that which we want to characterise formally, is given by the following of epistemic norms. It is then perspicuous to observe that a logician has *knowledge-how* of their logic when they have sufficient understanding of their practice and formalisation to make a sufficient range of valid or (with respect to the logic's semantics) correct inferences in it. Here it is appropriate to consider knowledge-how of a logic to be an ability, the ability to make a sufficient range of valid or correct inferences with respect to that logic. These logicians have *knowledge-that* of their propositions when they have rightly thought them to be the case with respect to their own theory of knowledge, that being a guide to their rational practice which their logic is a formal presentation of.[1] Knowledge-how is thus conceptually prior to knowledge-that.

Jason Stanley and Timothy Williamson in [73] have purported counterexamples to this relation of knowledge-how to ability, which ultimately rest on what for our purposes is an excessively broad understanding of what counts as an ability. For instance, a ski instructor might be said to have knowledge-how without ability if they have the "knowledge-how" of all the movements needed to make a complex manoeuvre, but no longer have the ability to execute it. Or a concert pianist might lose their arms in an accident, tragically losing the ability to play while still "knowing how" to do so. The response here is simply to distinguish knowledge-how of a logic from the kinds of knowledge-how that Stanley and Williamson invoke. This is not an ad hoc move, as under the present context of discussion we have already distinguished the relevant abilities as being particularly intellectual rather than physical, as they fit into our picture of rational practices.

Now let's return to the Propositional Thesis. A distinction should be drawn

[1] The words "rightly thought" are used rather than "truly believed", in order to remain agnostic on the question of whether knowledge-that reduces to justified true belief, or is taken to be primitive as in [89].

between following the norms underlying our practice and accepting them, the latter of which involves an evaluative judgment. We already understand that following epistemic norms is what we do in the process of engagement in our practice, which logic is meant to make explicit by formalisation. Accepting these norms implies an awareness of them and a commitment to their truth. But (contra Stanley and Williamson) we do not need such a commitment, or even such an awareness, in order to be engaged in a good practice. We might simply wish to be successful in what we do, or we might be open to amending our practice, or we might be engaged competently in a practice without knowledge of what we are doing.

In the domain of logical applications to scientific practice, for example in the formalisation of theories of physics, the sort of theoretical commitments that one makes turns out to be paradoxical, or at least contradictory, when the Propositional Thesis is assumed. Consider that, in having knowledge-how of a cutting edge scientific theory with respect to a particular domain, we ought to say we have the *right* or in some sense *best* theory of that domain. Under the Propositional Thesis, we ought to further say that what we can conclude from the theory's formalisation is (in particular) true. Yet by the pessimistic meta-induction, it is almost certain that some of our conclusions are false.

An analogous issue is the so-called *preface paradox*: after writing a monograph on a topic in which one is an expert, one has reason to be confident in the veracity of the book, but not on the truth of each individual constituent sentence of the book. Assuming that the veracity of the book relies on the truth of each of its constituent sentences, one has reason to be confident in the book not being true, in that one has reason to believe there are false sentences. One also has reason to be confident in the book being true, given that the writer is an expert on the subject.

Does the book show that the writer is an expert? The Propositional Thesis would have it that knowledge-how is knowledge-that of a set of propositions together with their consequences. The supposed expert is likely to then fall short in at least two ways. The first way is that the writer turns out to have made an error at some point in the monograph that contradicts a whole bunch of the consequences that would constitute knowledge-that. The second way is that the writer is elucidating a method that at some point in the future turns out to be out of date. The propositions that constitute knowledge-how have changed, perhaps even the means by which they are expressed. These ways in which the supposed expert would fall short both reveal a vicious skepticism that is manifestly unhelpful in evaluating claims of expert knowledge.

One might try to save the Propositional Thesis by referring to the constituent set of propositions for knowledge-how as merely being the right set of propositions with respect to a particular time. But what it is that makes the set of propositions the right choice would be something other than the propositions themselves, undermining the Propositional Thesis.

There is still another issue with the Propositional Thesis: epistemic norms are quite often not expressible by the sentences of the object language of the logic that they manifest through formalisation. Take for instance the "proof-theoretic norm of meaning", mentioned in an earlier section, that "if you want to know the meaning of a proposition A, you must have a means to prove or refute A, and if A is a compound formula then its proof (refutation) is explained in terms of the proof (refutation) of the relevant constituent subformulas", with an intuitionistic reading of "proof". The expression of this in a formal language would involve a quantification over sentences, which intuitionistic propositional and first-order predicate calculi (for instance) do not have the syntactic resources for. But in our conception of logical pluralism we wish to use both expressively stronger and weaker logics in accordance with our own desires. The Propositional Thesis might indeed sit better with logical monism, which we have already seen fit to dismiss on independent grounds, and would not help with our broader project for different practitioners to learn from one another.

It is necessary to draw a contrast between the above and Gilbert Ryle's influential argument against intellectualism in [66], that knowledge-how fails to reduce to, or is not a species of, knowledge-that on pain of regress:

> I argue that the prevailing doctrine leads to vicious regresses, and these in two directions. (1) If the intelligence exhibited in any act, practical or theoretical, is to be credited to the occurrence of some ulterior act of intelligently considering regulative propositions, no intelligent act, practical or theoretical, could ever begin... (2) If a deed, to be intelligent, has to be guided by the consideration of a regulative proposition, the gap between that consideration and the practical application of the regulation has to be bridged by some go-between process which cannot by the pre-supposed definition itself be an exercise of intelligence and cannot, by definition, be the resultant deed.

Stanley and Williamson respond to this by denying that the performance of actions entails the presence of knowledge-how to act, and by further denying that knowledge-that of a proposition itself involves an action of explicitly contemplating the proposition that one knows. It is worth stating a few apparent features of their characterisations of knowledge-how and knowledge-that: (1) for them knowledge-how is *intentional*, while not every action is; (2) for them knowledge-that does not require an act of contemplation, or any particular action whatsoever, but may be *implicit* in one's actions.

It is not necessary to attack or defend Ryle's argument here (the Propositional Thesis has itself been attacked enough, and there is not enough space to scrutinise his regress argument here), but rather it is more pertinent to further comment on Stanley and Williamson's characterisation of *knowledge-how as intentional* and *propositional knowledge as possibly implicit*. In the practice of logic as we have conceived it, neither characterisation is particularly helpful. Knowledge-how manifests itself as the ability to engage in rule-following for a rational practice, intentional or otherwise. In contrast, we know that something is the case when we have rightly thought it to be the case with respect to our own practice, which itself requires us to self-consciously follow the right sort of epistemic norms.

Here we have scrutinised a narrow version of the intellectualist account of knowledge-how, the Propositional Thesis, and seen that it is inadequate as a means of characterising the potential ability for the formal mathematician to make inferences, a characterisation that we would need for addressing the Problem of Communicability. It has already been mentioned that there are few formal accounts of knowledge-how, but there are some, which deserve further attention.

4.3 Lau and Wang on Epistemic Ability

Given the failure of the propositional thesis, and ongoing defences of "radical anti-intellectualism" such as [33], there have been remarkably few attempts at formalising knowledge-how. Among those that do exist are work by Yanjing Wang on *goal-directed knowledge how* ([86]), and Tszyuen Lau and Yanjing Wang ([43]) on *knowledge of an ability* important primers for what lies ahead. What follows is a brief treatment of Wang's goal-directed knowledge how, and Lau and Wang's knowledge of an ability.

First, the *goal-directed knowing how* predicate is defined by Wang as a binary modal predicate over a propositional calculus. Define the language $\mathcal{L}_{Kh}$ by $\phi \in \mathcal{L}_{Kh} ::== p|\top|\neg\phi|(\phi \wedge \phi)|Kh(\phi, \phi)$ where $Kh(\phi, \psi)$ is a predicate indicating that an agent knows how to do ϕ given ψ. The semantics of Kh is definable through a particular sort of labeled action model. In particular, consider a set of states S, a set Σ of action symbols, a set of *labeled actions* $R : \Sigma \to 2^{S \times S}$, and a valuation function $V : S \to 2^P$. We can speak of a labeled action in the following way: given two states $s, t \in S$, and an action a, we say that a is *executable* at s or $s \xrightarrow{a} t$ if $(s, t) \in R(a)$. Now if we consider sequences of actions Σ^*, assume that for the "empty sequence" $\epsilon \in \Sigma^*$ we always have $s \xrightarrow{\epsilon} s$. Furthermore, given a sequence of actions $\sigma = a_1...a_n \in \Sigma^*$ in general and $s, t \in S$, we say that $s \xrightarrow{\sigma} t$ if there exist $s_2, ...s_n \in S$ such that $s \xrightarrow{a_1} s_2... \xrightarrow{a_{n-1}} s_n \xrightarrow{a_n} t$. Define the initial segment of such a sequence σ by $\sigma_k = \epsilon$ whenever $k = 0$, and $\sigma_k = a_1...a_k$ whenever $n \geq k > 0$. We say such a sequence σ is *strongly executable* at a state s if: either $\sigma = \epsilon$, or $\sigma = a_1...a_n$

and for any k such that $0 \le k < n$ and any t, if $s \overset{\sigma_k}{\to} t$ then t must have at least one a_{k+1} successor.

This prepares us suitably for the semantics of L_{Kh}, in terms of what a model M provides relative to a state s:

- $M, s \vDash \top$;
- $M, s \vDash p \Leftrightarrow p \in V(S)$ for p an atomic sentence;
- $M, s \vDash \neg\phi \Leftrightarrow M, s \nvDash \phi$;
- $M, s \vDash \neg\phi \Leftrightarrow M, s \vDash \phi$ and $M, s \vDash \psi$;
- $M, s \vDash Kh(\phi, \psi) \Leftrightarrow$ there exists $\sigma \in \Sigma^*$ such that for all s' such that $M, s' \vDash \phi$: σ is strongly executable at s' and for all t such that $s' \overset{\sigma}{\to} t$, $M, t \vDash \psi$.

Lau and Wang's formalisation is here outlined because their conception of knowledge-how in terms of a successful path through states in an action model is a starting point for our approach. Yet for our purposes we aspire to an account of knowing-how that is what Wang refers to as *rule-directed* knowing-how, in particular that whose guidance is constituted by epistemic norms. The semantics of knowing-how should accordingly be different: in particular the assessment should be *local* to the agent's own information state. There will also have to be a lot more spelling out about how the actions must conform to the rules of the agent's own practice.

The work of Lau and Wang identifies a particular sort of propositional knowledge, *knowledge of an ability*. While their approach is a reduction of knowledge-how to knowledge-that, which (as we've seen from the discussion of the Propositional Thesis) won't suit our purposes, their pragmatic motivation for the intellectualist approach deserves a few words.

Lau and Wang distinguish between the *way* and the *method* of a practice, the way being an instance of the practice itself, the method being a description of that practice. Successful ways of a practice are stipulated conceptual givens in Lau and Wang's formalisation, while successful methods are what is expressed in the language. This allows them to clearly demarcate the boundaries of what they are trying to explain: merely a means of forming knowledge-how expressions rather than providing a meaning of knowledge-how itself, which it is "the very job of cognitive science" to provide through an account of various ways.

There is, in principle, more that the logician can do than this. If we can truly assert that we know how to achieve a task, it is just to say that we can follow the norms that must be followed, consciously or unconsciously, in order to achieve this task successfully. The guide to following these norms is just what logic formalises, so we can say that a knowledge-how predicate encodes the successful application of some logic. So far, one might respond that this conception of knowledge-how does not give us anything we can't have with a knowledge-that predicate following a conventional modal semantics, as we say that we know-how to achieve a deduction of ϕ exactly when we know that ϕ. Intellectualism thus returns. But logical omniscience prevents such an account from being remotely plausible; knowing how to use a logic does not involve knowledge of all the consequences. Rather, knowledge how involves knowledge of the appropriate formal procedure, in line with the relevant norms.

It will become imperative to outline what is meant by an "explicit reckoning of the procedures". This is obtained through relative interpretation of one logic inside another, insofar as that is possible; some examples will be given. But the grounds for doing this are given in a pragmatics of communication, intended to deal with the problem of communication for logical practitioners. Recent work by Teresa Kouri Kissel, already alluded to in chapter 1, is the starting point.

4.4 Kouri Kissel on Communicability

In chapter 1, it was mentioned that Teresa Kouri Kissel in [41] provides a means of characterising piecewise disagreement about the meaning of terms in terms of Craige Roberts' QUD framework. As already suggested, the implementation of QUD could not account for faultless disagreement in the cases where the logic shifts mid-conversation; this is elaborated more here.

Roberts' QUD framework, the context tuple $\langle I, G, G_{comm}, M, \leq, CG, QUD \rangle$ is roughly given with respect to a time t as follows:

- I is the set of interlocutors at t;
- G is a function from pairs (i, t), $i \in I$ to a *set of goals* such that $G(j, t)$ is j's set of goals at t;
- G_{comm} is the set of goals that all interlocutors in I have in common at t;
- $M = \langle A, Q, R, Acc \rangle$ is the set of *conversational moves* made by interlocutors up to t, with A the set of assertions, Q the set of questions, R the set of requests, and Acc the set of accepted moves;
- $\leq$ is the chronological ordering of all moves in the conversation;
- CG is the *common ground*, the set of presupposed propositions that all interlocutors share at t;

- QUD is the set of *questions under discussion*, the set of questions that are accepted, not entailed by any of the available presuppositions in CG ("available" in the sense of [72]), and whose answer is a goal all interlocutors have in common.[2]

Of particular note is the proposition CIP that Kouri Kissel invokes, the *correlation as identity proposition* indicating that the logical connectives are taken to mean the same thing. This may be viewed as a propositional form of interpretive charity: CIP is by default in CG, but not in cases where there is faultless disagreement about the meaning of logical terms. The following example that Kouri Kissel uses to illustrate this phenomenon is instructive:

> Rudolf: In my system, I can prove the fundamental theorem of the calculus. Does your system prove it?
> Geoffrey: I can prove a version of the fundamental theorem, too. I use nilsquares, numbers whose square is 0. They are not distinct from 0, but it is also the case that 0 is not the only one!
> Rudolf: What? Such nilsquares cannot exist. There is only one nilsquare and it is 0.
> Geoffrey: Well, using your negation, sure.
> Rudolf: Oh, I see. Your negation must behave differently from mine.

The interlocutors are Rudolf and Geoffrey, and there six distinct time frames.[3] At t_0, before the first sentence is uttered, CG includes all the axioms of a system of classical analysis, along with CIP. Over the course of the conversation it is clear that Rudolf is not sufficiently familiar with smooth infinitesimal analysis, so axioms of a system of smooth infinitesimal analysis are not in CG. At t_1, Rudolf adds the proposition that his system proves the fundamental theorem in CG, A, and Acc, and adds the question of whether the other system proves the fundamental theorem to Q and QUD. At t_2 Geoffrey answers that question and proposes to add the information that his system uses "nilsquares ... not distinct from 0, but it is also the case that 0 is not the only one" to CG and A, but at t_3 Rudolf tries to block this from CG and Acc and "[Non-trivial] nilsquares cannot exist" is added to A. At t_4 Geoffrey promotes the information that his negation is different to Rudolf's as an available implication in CG and adds the corresponding assertion to A. For the sake of charity, CIP is removed from CG. At t_5 propositions about the differences in connective meaning and the relative existence of nilsquares are added to CG, along with Geoffrey's assertions ad the proposition that smooth infinitesimal analysis proves a fundamental theorem.

[2]The stipulation that the presuppositions in CG must be "available", that is to say easily inferable, is Kouri Kissel's original contribution. Otherwise many potential questions under discussion would be prematurely ruled out through logical omniscience.

[3]This is a short paraphrase of Kouri Kissel's own explanation.

The upshot of the above analysis is that in the above particular conversational context, the meaning of negation turns out to be ambiguous between the kind that is used by Geoffrey and the kind that is used by Rudolf. Though the ambiguity is only realised by Geoffrey and Rudolf through a mutually charitable reading of the other participant's interpretation of the connectives, negation isn't ambiguous in itself. After all, the meaning of negation in itself is an external question, by Kouri Kissel's reading of Carnap.

A lot of the groundwork for a framework of understanding and agreement is nicely laid out for us, but there are still two significant kinds of limitation. The first kind of limitation is that the QUD framework cannot account for all kinds of faultless disagreement, and this is perhaps suggestive of a problem with the demand to exclude "external questions" about the meaning of the connectives.

To see the first kind of limitation in practice, it is worth applying the QUD framework again. Take for instance the following conversation.

> Arend: My propositional calculus can be expressed using only one operator, how about yours?
> Henry: Mine can indeed, and I can also define that operator in terms of only negation and disjunction.
> Arend: I can't do this: I need conjunction, disjunction, negation, and implication.
> Henry: I suppose you would need to have the law of excluded middle to do any better.

As with Geoffrey and Rudolf, in order for Arend to accommodate Henry's statements, Arend must allow for a *context shift*. In order for Arend to understand Henry's reference to "negation" and "disjunction", CIP must be removed from CG, and given the conversational context it's clear it wasn't there to begin with. Yet in order for Arend to understand Henry's reference to "the law of excluded middle", CIP must be added back to CG; the "law of excluded middle" as Henry puts it is understood by the classical reading of the connectives. In any case, the second and third statements are essentially about the operator Henry and Arend have defined, their being understood by the other party necessitates CIP staying out of CG, and as the fourth statement refers to the third, the fourth statement in turn requires CIP to stay out of CG. We reach an impasse.

The general point is that there are infinitely many conversations one can have about a hypothetical operator and its properties which would considerably obscure the playing field of what constitutes *the* context of discussion. A quick fix would be to rework the QUD framework to allow multiple contexts to coincide in the interpretation of a single sentence, but this would already take us quite far from the spirit of Kouri Kissel's suggestion: that the constraints on the meaning of the logical connectives should be solely given in the context of discussion. More importantly, amending the QUD framework in such a way would be an ad hoc move, without a view to what sort of principles ought to underlie the coinciding of multiple contexts.

The second limitation is that it is not clear on what grounds the subjects of conversation can adequately interpret one another. Here there are a number of points to make. An underlying feature of the conversation, that allows the pragmatics to work after the fact, is that Geoffrey knows to make the pertinent point about the negation available to Rudolf. But on what grounds does Geoffrey interpret Rudolf's use of negation? The axioms of a system of classical analysis are in CG, but Geoffrey can't simply add them to the axioms he works with (on pain of triviality). Geoffrey could interpret smooth infinitesimal analysis inside a classical metatheory, in which Rudolf's use of negation is interpreted. But then it wouldn't be clear that Geoffrey was truly *doing* smooth infinitesimal analysis, being open to accept Brouwerian counterexamples on their face.

Kouri Kissel invokes a propositional principle of charity for Rudolf to accept Geoffrey's revision of the logical connectives. This is insufficient for two reasons. First, Rudolf and Geoffrey are operating in different languages, and the scope of their charity is thus limited in different ways. In particular, on the face of it, Rudolf is not able on charitable grounds to interpret Geoffrey's assertion about nilsquares as true, because the maximally true family of beliefs that Rudolf is obliged to attribute to Geoffrey is a set of propositions with Rudolf's own connectives including his negation. Kouri Kissel is aware of this (my emphasis):

> Intuitively, if we assumed that both participants would interpret themselves as using the same connectives, we would have a problem with charity and accommodation. Geoffrey has uttered something that can be taken to imply the negation of a classical truth. His utterance can be taken to imply that there exists some non-singleton, non-empty, set of things that are not distinct from 0. Since the double negation cancels in classical logic, this is equivalent to saying that these things both are and are not distinct from 0 for the classical analyst–a clear contradiction. Adding the existence of these types of nilsquares to a classical system will result in a contradiction. Rudolf cannot accept the existence of these objects and maintain a consistent theory. **If the conversation continues after the existence of such nilsquares is posited (and not just by a simple 'no' from Rudolf), then we must assume that Rudolf is assuming that Geoffrey is using a different negation.** So, the most charitable interpretation of this conversation is one in which the negation connectives have

different meanings.

The highlighted passage makes it clear that the account of how Rudolf accepts the revision is made after the fact, there isn't a principled reason behind Rudolf accepting Geoffrey's move to assert that his negation is different. The "charitable interpretation of this conversation" invoked is that of charity on behalf of the reader interpreting the transcript, not Rudolf interpreting Geoffrey.

The second problem with the use of the propositional principle of charity is that it does not provide an account of how Rudolf comes to understand Geoffrey, not just accept Geoffrey's remarks as faintly coherent. The circumstances under which Rudolf, who knows himself to be an expert, should treat Geoffrey's remarks charitably are a complex sociological matter. But a necessary condition is that Rudolf should be able to understand what Geoffrey says, even if he doesn't accept it, but being limited by his prior interpretations of the logical connectives it's not clear how he could.

A curious feature of the family of examples exemplified by Rudolf and Geoffrey is that the common ground has not been sufficiently updated because one party or another lacks sufficient *propositional knowledge* of what the other knows. In particular, from the start Rudolf does not know the axioms of smooth infinitesimal analysis. In general, this does not appear to be a wise starting point for a productive conversation. Rudolf should know how to prove basic facts about smooth infinitesimal analysis to understand Geoffrey's remarks and adequately assess whether he and Geoffrey are in a state of faultless disagreement.

The alternative pursued here departs instead from the notion of "common ground" altogether, and the propositional variety of the principle of charity being used, but instead makes use of the practice-oriented principle of charity to motivate relative interpretation. What follows is an ostensive account of knowledge-how in terms of labeled proofs, and show how that together with the practice-oriented principle of charity makes for a more accommodating pragmatics of communication between practitioners of logic.

4.5 The Present Approach

4.5.1 Knowledge-How

Here an ostensive notion of knowing-how is introduced. Being situated in a pragmatics of communication between practitioners of logic, it is appropriate to start with conditions of assertion:

- For a logician to assert a proposition ϕ, she ought to be able to outline its proof (with respect to their own practice and information state);

- Her having the ability to outline a proof of ϕ means she knows how to follow her practice from her prior information state to ϕ.

At this point we still haven't provided a substantive account of what it is that knowledge-how gets us, particularly over and above propositional knowledge. What follows is a more detailed heuristic account of what knowledge-how is, with reference to Wang's previous work. We consider a subject s to have at time t a *labelled inference* $P_{s,t,\phi} = \langle S, \Sigma, Act, LAct, V \rangle$. Here S is a set of states, Σ is a set of *histories*, ordered sets of states. The set Act is the set of action symbols, $LAct$ is the set of *labelled actions* $LAct : Act \to \bigcup_{i \in I, j \in J} 2^{S_i \times S_j}$ that say whether $a \in Act$ sends state S_i to S_j, and a valuation function $V : S \to 2^Q$ where Q is the set of propositions ψ contained in some $\alpha \in S$.

In our account of knowledge-how, it will be necessary to provide a notion of validity that relativises validity according to a "weakly faithful" translation, which must first be explained. Given two subjects s_1 and s_2 each with their own logic, a *translation* $E(s_2, s_1)$ of s_2 according to s_1 is a map from the formulas of $\mathcal{L}_{s_2}$ with only arithmetical variables free to the formulas of $\mathcal{L}_{s_1}$ such that for every formula ϕ of $\mathcal{L}_{s_2}$ and its translation $\phi^{E(s_2,s_1)}$, we have $FV(\phi^{E(s_2,s_1)}) \subseteq FV(\phi)$. Moreover, we say that $E(s_2, s_1)$ is *weakly faithful* when for any set of $\mathcal{L}_{s_2}$-formulas Φ such that $\Phi \nvDash_{s_2} \bot$ there is a set of $\mathcal{L}_{s_1}$-formulas Ω such that $\Omega \nvDash_{s_1} \bot$ and for any $\psi \in \mathcal{L}_{s_2}$, we have $\Phi \vDash_{s_2} \psi$ iff $\Omega \vDash_{s_1} \psi^{E(s_2,s_1)}$.

A weakly faithful translation doesn't have to always exist. But it is a suitable condition for making use of results worked out in one logic inside of another. In the event that such a translation exists, we are now in a position to characterise relative validity in terms of that translation. Let $E(s_2, s_1)$ be a weakly faithful translation of s_2's logic into s_1's. A labelled inference $P_{s_1,t,\phi} = \langle S, Act, R, V \rangle$ is said to be $E(s_2, s_1)$*-valid* when (1) for all ϕ, ψ such that $\phi \leq \psi$, then $\phi \vDash_{s_1} \psi$, and (2) the translation $E(s_2, s_1)$ of s_1's logic into s_2 is weakly faithful.

Consider two subjects s_1, s_2 at a time t. We say that s_1 has a *labelled proof* $P_{s_1,t,\phi}$ of ϕ at t when the labelled inference $P_{s_1,t,\phi} = \langle S, Act, LAct, V \rangle$ of ϕ at t is $E(s_2, s_1)$-valid.

It should be emphasised that there is, under this conception, no unique choice of labelled proof that a subject is meant to have. The choice is only bounded by the condition that the proof is sound with respect to the translation of the interlocutor's logic and assumptions. But questions can be used by the interlocutor to prompt the subject to expand any single step of the proof into more valid steps. Recall our standard of assertion; followed through, a logician should in principle be able to *reason coinductively* from a valid proof in broad sketches to a proof in minute ones that each correspond to, for instance, some invocation of an axiom of rule of inference.

So we say that a logician s_1 can be said by an interlocutor s_2 to have knowledge-how of a proof of ϕ, as a condition of being able to assert ϕ, when:

1. s has a labelled proof $P_{s_1,t,\phi}$ of ϕ at t;

2. When at $t+n$ for some nonnegative integer n, s_1 is asked whether ϕ, s_1 can provide a labelled proof $P_{s_1,t+n+1,\psi}$ of ϕ with respect to Σ.

So an account of knowledge-how, together with an adequate account of translation between formal languages, allows a subject to communicate parts of a proof, not necessarily the entirety, in order to fill in missing pieces of information for the interlocutor. But how knowledge-how is particularly used is best seen through an updated framework for a pragmatics of communication between practitioners of logic.

4.5.2 Returning to the Pragmatics of Interaction

We outline the pragmatics of a disagreement with Craige Roberts' question under discussion framework as a starting point ([62]). This is the same starting point as that of Teresa Kouri Kissel's pragmatic framework for logical instrumentalism in [41], which we have already criticised, but there is one crucial difference. Instead of (as in Kouri Kissel's proposal) a "common ground" of propositions that interlocutors share, and a principle of charity that maximises common ground, there is a mutual understanding of each party's epistemic norms with respect to a principle of charity that maximises the norms' utility. Given subjects a and b, the context tuple $\langle I, G, G_{comm}, M, \leq, E, QUD\rangle$ of their discussion is roughly given with respect to a time t as follows:

- I is the set of interlocutors at t;

- G is a function from pairs (i,t), $i \in I$ to a *set of goals* such that $G(j,t)$ is j's set of goals at t;

- G_{comm} is the set of goals that all interlocutors in I have in common at t;

- $M = \langle A, Q, R, Acc\rangle$ is the set of *conversational moves* made by interlocutors up to t, with A the set of assertions, Q the set of questions, R the set of requests, and Acc the set of accepted moves;

- $\leq$ is the chronological ordering of all moves in the conversation;

- $E = \{E_{x,y} | x, y \in \{a,b\}\}$ where $E_{x,y}$ is a weakly faithful *translation* conventionally taken of the logic of an interlocutor x according to a subject y;

- *QUD* is the set of *questions under discussion*, the set of questions that are accepted, not entailed by any of the available consequences of $E_{a,b}$ and $E_{b,a}$, and whose answer is a goal all interlocutors have in common.[4]

The questions may be *conclusively agreed* if for both parties a statement has been proven or refuted, or *conclusively disagreed* if the parties reach different conclusions, but in other cases they may draw on their knowledge of the other parties' knowledge-how to inform their own proofs. Some questions may lead to *inconclusive* results in the sense that the conversations may never result in either sort of conclusion.

Much of the following will be given in a setting where the agent is a classical logician and the interlocutor is an intuitionist, as this is a relatively familiar setting in order to discuss further, heretofore unmentioned complications. For instance, in the forthcoming language the agent and the interlocutor are not interchangeable, as metalogic is not neutral. We discuss some different conclusions that might arise from, say, having the agent Int be an intuitionist and the interlocutor Class be a classical logician, trying to work on a problem together.

4.5.3 A Case Study

An account of just how Int and Class interpret each other would, under present understanding, be ad hoc in certain respects. There is no single standard approach to translating intuitionistic mathematics in classical mathematics, or vice versa. For the former, we will employ the Henkin interpretation as it is used by Albert Visser in [85]; for Int that can talk about theorems that Class knows how to prove. For the latter, we will employ Gödel's double negation translation.

In light of the ad hoc nature of the interpretation, one might wonder what the purpose of the ensuing discussion is. The purpose is to illustrate one means to obtain an explanation of the following, as in the conversation examples given in the previous subsection: (1) how agreement is expressed; (2) how disagreement is expressed; (3) how "higher-level disagreement" leads to inconclusive results. Whether another helpful choice of interpretation of classical mathematics into intuitionistic mathematics and vice versa would produce dramatically different outcomes in this regard is doubtful.

The forthcoming will be from the point of view of Class. Intuitionistic arithmetic can be translated inside of classical arithmetic using the Henkin construction: one can define in classical arithmetic a set of $\mathcal{L}$-sentences Δ that is saturated,

[4]That is "available" in the sense of [72], roughly "easily inferable", as Kouri Kissel also refers to; this is only a slight modification of Kouri Kissel's own original contribution, as she starts from common ground rather than $E_{a,b}$.

and iterating this process can provide a Henkin construction of a Kripke model of intuitionistic arithmetic.

Now for the point of view of Int. There are a few differing means of translation from classical arithmetics into intuitionistic arithmetics; Albert Visser in [85] provides a characterisation of several different ones. For the sake of clarity we'll stick with one: Gödel's double negation translation.

Gödel's double negation translation simply maps every formula A of PA into a formula A' which is syntactically the replacement of every subformula of A that is a disjunction of an existential quantification by its double negation. It is already known that the double negation translation is weakly faithful, as it preserves $\bot$ and the double negation of every theorem of PA is a theorem of HA.[5]

The above translations also relevantly apply to extensions of PA and HA that add exponential and logarithmic functions, as in some of the below examples.

Three examples we claim to account for are given below. The first concerns a statement with a short proof in classical mathematics, and a more difficult proof in intuitionistic mathematics:

> Class: I can prove there exist real numbers a and b such that a and b are irrational, but a^b is rational.
> Int: How do you know?
> Class: I know $\sqrt{2}$ is irrational; $\sqrt{2}^{\sqrt{2}}$ is either rational or irrational .
> . .
> Int: How do you know $\sqrt{2}^{\sqrt{2}}$ is either rational or irrational?
> Class: From the law of excluded middle.
> Int: I know how you can prove the theorem, but I can't.
> Class: Can you prove that $\sqrt{2}$ is irrational?
> Int: Yes. I can also prove that $2log3/log2$ is irrational.
> Class: Do you know how to prove the theorem now?
> Int: Yes, as $\sqrt{2}^{2log3/log2} = 2$.
> Class: I have another proof of this theorem.

Interlocutors Class and Int agree on the question under discussion and particularly the goal: to prove the theorem in question. They each understand the epistemic norms that constitute the other's practice. The unfolding of knowledge-how allows Class and Int to present steps of a proof without spelling out all of

[5] On the other hand, the translation does not preserve disjunction. Furthermore, Visser has proven that if any translation of PA to HA is weakly faithful and preserves modus ponens, as the double negation translation is and does, it cannot preserve disjunction.

the steps, as the practice-based principle of charity allows each interlocutor to fill these in. The practice-based principle of charity allows Int to interpret Class' non-constructive proof, but asks further questions to confirm that Class' proof is indeed non-constructive. Moreover, constructive proofs that $\sqrt{2}$ and $2log3/log2$ are irrational are also non-constructively valid, as Class should realise from a comparison of their respective epistemic norms.

At t_0, before the first sentence is uttered, the theorems of $E_{Int,Class}$ include the axioms of intuitionistic logic embedded in extensions of PA by the aforementioned Henkin construction; the theorems mapped under $E_{Class,Int}$ include double negation interpretation equivalents of the axioms of classical arithmetic inside of intuitionistic arithmetic. At t_1, Class asserts that there exist real numbers a and b such that a and b are irrational, but a^b is rational. At t_2 Int asks Class to support his assertion, which Class proceeds to outline in a classical manner. At t_6 Int admits that he lacks a proof of the theorem by his own lights. Class understands from the Henkin construction that there is a proof, and proceeds to make requests of Int, until Int has his proof.

A different kind of example leads to inconclusive results:

> Class: How would you evaluate the function $g : \mathbb{N} \to \mathbb{N}$ such that $g(x) = 1$ if the Goldbach conjecture is true, and $g(x) = 0$ otherwise?
> Int: I take it that g is not well-defined, as there is no proof the Goldbach conjecture is true.
> Class: How would you evaluate the function $h : \mathbb{N} \to \mathbb{N}$ such that $h(x) = 1$ if given that the Goldbach conjecture is true, the Goldbach conjecture is true, and $h(x) = 0$ otherwise?
> Int: I take it that h is not well-defined, as there is no proof the Goldbach conjecture is true, so there is no proof that given that the Goldbach conjecture is true, the Goldbach conjecture is true.
> Class: If there was a constructive proof of the Goldbach conjecture, g and h would be well-defined and continuous for both of us.
> Int: There is no proof of the Goldbach conjecture.

What is it about the last conversation that leads to inconclusive results? There is *higher-level disagreement*, at the level of the respective metatheories of the interlocutors. In this case, Int cannot agree to entertain a hypothesis without the support of a constructive proof. Interpretive charity can't treat this sort of disagreement when it is performed with respect to the practice of the interlocutors as they manifest in object-level propositions, but an analogous charity principle at a higher level could.

That is to say at t_0, Class and Int interpret one another through $E_{Int,Class}$ and $E_{Class,Int}$ respectively, but the question at t_1 and t_3 concerns the input of a function, which Int and Class must differ on in principle. There is no resolution unless and until the Goldbach conjecture is solved, which would allow Int to accept Class' asserted conditional and evaluate the function.

Finally, we return to the example of Arend and Henry.

> Arend: My propositional calculus can be expressed using only one operator, how about yours?
> Henry: Mine can indeed, and I can also define that operator in terms of only negation and disjunction.
> Arend: I can't do this: I need conjunction, disjunction, negation, and implication.
> Henry: I suppose you would need to have the law of excluded middle to do any better.

At t_0, Arend and Henry interpret one another through $E_{Arend,Henry}$ and $E_{Henry,Arend}$ respectively, and what follows is a series of assertions. In order for Henry to assert "you would need to have the law of excluded middle" he has an understanding of Arend's interpretation of excluded middle and its implications for the relative definability of certain connectives.

We've seen, then, that the new approach can accommodate a range of cases of agreement that Kouri Kissel's fails to account for, and also some cases of faultless disagreement as well which make explicit certain tensions in trying to think of classical logic intuitionistically, or think of intuitionistic logic classically.

4.6 Conclusion

In an earlier section, Teresa Kouri-Kissel's pragmatic account of faultless disagreement between practitioners of different logics was treated. There were three issues with the treatment, which otherwise worked as a starting point: first, there is fault in principle with the view that the appropriateness of logics in and of themselves should be treated as Carnap would an "external question"; second, the account does not provide for context-shifts in the understanding of which logic is under discussion that would be pervasive in such situations; third, the account cannot show a positive interpretation of the interlocutor's beliefs, only an awareness (though Kouri-Kissel's own propositional principle of charity) that the meanings of the operators must be different from the interlocutor, which doesn't appear to be what happens (or even just what should happen) when different practitioners meet. The first issue is accounted for in chapter two, where it was argued that logic provides us with evidence that may compel us to revise our norms on the basis of a narrow reflective equilibrium. Furthermore, differences between practitioners may be manifestations of differences in ethics or strategy; interrogations of these differences are not all meaningless "external questions", but that topic is too far from the subject of this dissertation to discuss further. The second and third issues have been accounted for here. The change in orientation *from* beliefs and assigned meaning–*to* fragments of language to knowledge-how and assigned strategy to the class of possible proofs–provides a basis for mutual understanding rather than simply understanding of difference.

There is an underlying anxiety about the normativity of logic that this dissertation could at least partially address. The anxiety is roughly that logic: (1) being the form of all reasoning as such, is of only one kind, so it's possible none of us have gotten to (or could ever get to) "reason logically" at all; (2) it is unclear in what sense logic does or could change its own rules; (3) there is no a priori account of how vicious semantic paradoxes can be reliably identified at every level; (4) there is no a priori account of how practitioners of different logics understand one another. In what has preceded, hopefully this multifaceted anxiety has been resolved.

Chapter 5

Conclusion

Here we approach the end of the dissertation, but by no means have we approached the end of the subject. More needs to be said about future directions. The first part of this conclusion summarises the lessons of the previous chapters. The second and third parts provide suggestions for future work.

5.1 The Lessons of the Previous Chapters

The first chapter provided an account of Thoroughgoing Logical Pluralism, appropriately so called as it is founded on a pluralism about epistemic value with respect to what constitutes good practice in thought and behaviour. This contrasts with accounts of logical pluralism emerging from linguistic pluralism attributed to Carnap, the logical nihilism of (for instance) Gillian Russell, the pluralism-of-Tarskian-cases approach of Beall and Restall, or Pedersen's logical pluralism founded on Lynch's alethic pluralism. Motivation for this position is provided by the response to the collapse argument against logical pluralism and the groundwork for communication between practitioners on the basis of the practice-oriented principle of charity that avoids the charge of solipsism that Davidson's version of the principle of charity (for instance) may face.

The problem of how a single practitioner reflects on their own practice, or their own logic, on the basis of logical consequence is addressed in the second chapter. The short answer is that they do so by means of pursuing narrow reflective equilibrium of the norms that constitute their own practice with the rules and axioms, or the mathematical structures characteristic of the class of models, of the logic in question. This is by contrast with both "inference to the best explanation" approaches (as characterised by the recent work of logical anti-exceptionalists) and the pursuit of wide reflective equilibrium.

The prior work was inspired by how it makes sense to think of the semantic paradoxes as inspiring revision in some logical principles, or not as the case may be. This would depend on the theoretician's priorities, but would still be determined by narrow reflective equilibrium. There is nothing we can say for sure about the semantic paradoxes that distinguishes them as more than a particular artefact of the (logical) language used to formulate them, which can be predicted and detected using Kripkean fixed point methods. In the third chapter it was established that even assuming that liar sentences are prima facie true and false, along with all the "revenge" paradoxes that arise from regimentation of liar paradoxes from truths and falsehoods, there are ways to account for them all the way up. The predictability of the semantic paradoxes, and the narrowness of their application, allows practitioners to bracket them out from most of their practice. The semantic paradoxes may be expressed if such an expression serves a purpose, for instance to account for some other paradoxes like those of set theory.

Another way to think of the upshot of the first and second chapters is that logic provides a practitioner with grounds to reflect on whether their practice is adequately conceived, and whether their logic is a good fit for their practice. The third chapter then provides a case study of whether the logic is indeed a good fit for their practice in terms of an evaluation of whether the linguistic resources employed are necessary. What is still missing is a more thorough account of the pragmatics of how one practitioner may evaluate another in terms of whether they substantively disagree, already suggested in chapter one but not elaborated further. This is the subject of chapter four, where knowledge of what can be learned from another practitioner arises from knowledge-how of the other's practice. Such a perspective is incompatible with intellectualist accounts of knowledge-how.

In what remains I discuss an overarching limitation of the approach, in the conception of practices being constituted by epistemic norms, and the potential of the account to provide an answer to the problem of the normativity of logic.

5.2 Beyond Descriptive Languages

When logic is characterised as a formal presentation of a guide to following a rational practice, a practice itself constituted by epistemic norms, this applies solely to *descriptive* languages. At least, *inquisitive* statements like "Do I stay at home?" do not in general provide us with a guide to developing knowledge, but rather proposals that an agent can respond to in different ways.[1] And *normative* statements like "You ought to stay at home" provide constraints on agents rather than a guide. This would appear to be a profound limitation to the present ac-

[1]For instance, Ivano Ciardelli and Floris Roelofsen write in [13] that inquisitive semantics "views propositions as proposals to update the common ground".

count of logic.

Thankfully, there is reason to believe that the above account can be generalised according to the goals of the practitioner. But as those goals may be socially oriented, the account of what logic would be for the various practitioners would become complicated very quickly, and a topic all on its own. Below is an account of what is meant by this.

Logics in normative and inquisitive languages are still *formal presentations* as already conceived. They provide instructions to follow some rules. But “You ought to stay at home” is properly conceived as a consequence of constraints on an agent’s possible goals. “Do I stay at home?” is an instance of a class of proposals to entertain an agent’s possible goals. This invites the following generalisation: a logic is a formal presentation of a guide to following rules that either constrain, entertain, or further an agent’s possible goals.

The “or” in “either constrain, entertain, or further an agent’s possible goals” is not exclusive. After all, many inquisitive and normative logics are *mixed* formal languages, with something like a question operator or an “ought” operator in an otherwise descriptive formal language. There might also be multiple agents, all affected by (say) a public announcement.

The more interesting consequences of this more general conception come from trying to find analogues for the notion of logical revision and the pragmatics of communication between practitioners explored later in the thesis. At a minimum, to find analogues for much of what was discussed in those chapters one should aim to protect the basic commitments to: (1) a practice-oriented principle of charity; (2) logical revision being on the basis of seeking narrow reflective equilibrium between the consequences of the logic and (more generally) the rules that either constrain, entertain, or further the relevant agent’s possible goals. A great deal of theoretical questions are left open, for instance:

1. When should logical revision takes place? Where there is communication between different practitioners working on different logics, there may be cause for one agent to revise their logic after another has done so themselves.

2. How do we define the totality of possible goals for an agent?

3. Say there are rules meant to further an agent’s possible goals and other rules meant to constrain them. Are they assigned the same sort of weight for the purposes of logical revision?

5.3 The Normativity of Logic

A central tenet of the present account is that logic is normative for thought insofar as it outruns the mental capacities of practitioners to think through the consequences of their norm-following. This is a generalisation of a couple of different insights with a strong upshot worth picking apart in future work.

The first insight is an account by Catarina Dutilh Novaes in [23] into the *operative role* of formal languages for practitioners, that rendering thoughts into formal languages provides a heuristic function for the logician akin to writing on a blackboard for a mathematician. Formal expressions might clarify inferences in their mode of reasoning that would not be available if the propositions were not expressed in this particular way. Furthermore, they serve to make explicit the differences between the reasoning we wish to encode and the erroneous inferences we may still instinctively take, for which she discusses some relevant findings on recurrent pathologies in reasoning that might have developed as an adaptation.

The second insight is the increasing body of work by mathematicians and computer scientists into the development of automated theorem proving for fragments of mathematical language. Certain proofs of long-standing mathematical problems (the four color theorem, the Robbins conjecture, the Kepler conjecture) were made with automated theorem provers.[2]

This is not to say that the use of formal languages is closely tied to the possibility of their implementation in automated theorem proving. Even where logics and formal languages are not easily implementable, formal languages are still useful in general for their operative role that Dutilh Novaes outlines. The present account differs from Dutilh Novaes in two ways. First, all logic, so conceived, is formal. Second, the account here does not merely spell out that formal languages serve a heuristic function for the practitioner, but that logics (always being formal) generate their own consequences, with the writing down of expressions and the implementation of proof assistants as being ways for the practitioner to gain privileged access to them. The "operative role" is the result of the practitioner's access, but the implementation of logics into automated theorem provers are other ways for a practitioner to gain privileged access to perhaps unforeseen consequences. The two insights are instances of the same overarching quality of logic, that it outruns our capacity to think through the consequences of our norm-following.

For descriptive languages, we have already seen that logical revision is a form

[2] A survey of results, already out of date but still concerning those results among many others, is found in [31].

of narrow reflective equilibrium on the logic's consequences and the epistemic norms constituting the practice for which the logic is a guide. So given the already-considered utility of the potential of logics to outrun the mental capacities of practitioners to think through the consequences of their norm-following, considerations of parsimony lead us to a simple account of the normativity of logic: logic is normative for thought insofar as it outruns those capacities. There is no other reason that one ought to follow logical consequence in particular than this.

For languages that are not descriptive, an account of the normativity of logic may be given in terms of goals rather than epistemic norms, but as already suggested in the prior section the complexities of such an account exceed the scope of the thesis.

Bibliography

[1] Anderson, A. R., Belnap, Nuel, and Dunn, J. M., *Entailment: the Logic of Relevance and Necessity* Vol. 2, Princeton University Press, 1992.

[2] Artemov, Sergei and Protopopescu, Tudor, "Intuitionistic Epistemic Logic" *The Review of Symbolic Logic*, Vol. 9, No. 2 pp. 266–298, 2016.

[3] Avigad, Jeremy, Donnelly, Kevin, Gray, David and Raff, Paul, "A Formally Verified Proof of the Prime Number Theorem" *ACM Transactions on Computational Logic*, Vol. 9, Issue 1, Article 2, December 2007.

[4] Azzouni, Jody, *Talking About Nothing: Numbers, Hallucinations, and Fictions*, Oxford University Press, 2010.

[5] Beall, J. C., and Restall, Greg, "Logical Pluralism", *Australasian Journal of Philosophy* vol. 78, pp. 475–493, 2000.

[6] Beall, J. C. and Restall, Greg, *Logical Pluralism*, Oxford University Press, 2006.

[7] Belnap, Nuel, and Green, Mitchell, "Indeterminism and the Thin Red Line", *Philosophical Perspectives 8*, James E. Tomberlin (ed.), pp. 365–388, 1994.

[8] Belnap, Nuel, and Gupta, Anil, *The Revision Theory of Truth*, MIT Press, 1993.

[9] Bishop, Errett, *Foundations of Constructive Analysis*, McGraw-Hill, 1967.

[10] Brandom, Robert, *Making it Explicit: Reasoning, Representing, and Discursive Commitment*, Harvard University Press, 1994.

[11] Brandom, Robert, *Articulating Reasons: An Introduction to Inferentialism*, Harvard University Press, 2001.

[12] Brouwer, L. E. J., "On the Foundations of Mathematics", In Arend Heyting (ed.), *L. E. J. Brouwer Collected Works 1: Philosophy and Foundations of Mathematics*, North-Holland, 1975.

[13] Ciardelli, Ivano and Roelofsen, Floris, "Inqusitive Logic" *Journal of Philosophical Logic* Vol. 40, pp. 55–94, 2011.

[14] Cardelli, Ivano A. and Roelofsen, Floris, "Inquisitive Dynamic Epistemic Logic", *Synthese*, Vol. 192, pp. 1643–1687, 2015.

[15] Carnap, Rudolf, *The Logical Syntax of Language*, trans. Amethe Smeaton, New York: Harcourt, Brace and Company, 1937.

[16] Carroll, Lewis, "What the Tortoise Said to Achilles", *Mind*, Vol. IV, Issue 14, April 1895.

[17] Cohnitz, Daniel, & Estrada-González, Luis, *An Introduction to the Philosophy of Logic*, Cambridge University Press, 2019.

[18] Coquand, Thierry, and Huet, Gérard, "The Calculus of Constructions", *Information and Computation*, Vol. 76, pp. 95–120, 1988.

[19] Cook, Roy T., "What is a truth value and how many are there?", *Studia Logica* Vol. 92 No. 2, pp. 183–201, 2009.

[20] Davidson, Donald, "On the Very Idea of a Conceptual Scheme", *Proceedings and Addresses of the American Philosophical Association*, Vol. 47, pp. 5–20, 1973.

[21] Dummett, Michael, *Thought and Reality*, Oxford University Press, 2006.

[22] Dutilh Novaes, Catarina, "The Different Ways in which Logic is (said to be) Formal", *History and Philosophy of Logic*, Vol. 32, Issue 4, 2011.

[23] Dutilh Novaes, Catarina "Towards a practice-based philosophy of logic: formal languages as a case study" *Philosophia Scientiae* Vol. 16, No. 1, pp. 71–102, 2012.

[24] Dutilh Novaes, Catarina, "A Dialogical, Multi-Agent Account of the Normativity of Logic", *Dialectica*, Vol. 69, Issue 4, pp. 587–609, December 2015.

[25] Field, Hartry, *Saving Truth from Paradox*, Oxford University Press, 2008.

[26] Field, Hartry, "Pluralism in Logic", *The Review of Symbolic Logic*, Vol. 2, No. 2, pp. 342–359, June 2009.

[27] Gonthier, Georges, "Formal Proof – The Four Color Theorem", *Notices of the American Mathematical Society* Vol. 55, No. 11 pp. 1382–1393, December 2008.

[28] Goodman, Nelson, *Fact, Fiction, and Forecast*, Fourth Edition, Harvard University Press, 1983.

[29] Harman, Gilbert, *Change in View: Principles of Reasoning*, MIT Press, 1986.

[30] Harman, Gilbert & Kulkarni, Sanjeev, *Reliable Reasoning: Induction and Statistical Learning Theory*, MIT Press, 2007.

[31] Harrison, John, Urban, Josef and Wiedijk, Freek, "History of Interactive Theorem Proving", *Handbook of the History of Logic* Vol. 9, pp. 135–214, 2014.

[32] Herzberger, Hans G., "Notes on Naive Semantics", *Journal of Philosophical Logic* Vol. 11, No. 1, pp. 61–102, February 1982.

[33] Hetherington, Stephen, "How to Know (that Knowledge-that is Knowledge-how)", Epistemology Futures pp. 71–94, S. Hetherington (ed.), Oxford University Press, 2006.

[34] Heyting, Arend, *Intuitionism: An Introduction*, third edition, North-Holland Publishing Company, 1971.

[35] Hjortland, Ole Thomassen, "Anti-exceptionalism about logic", *Philosophical Studies*, Vol. 174, pp. 631–658, 2017.

[36] Horsten, Leon, *The Tarskian Turn: Deflationism and Axiomatic Truth*, MIT Press, 2011.

[37] Kennedy, Juliette, "On Formalism Freeness: Implementing Gödel's 1946 Princeton Bicentennial Lecture", *The Bulletin of Symbolic Logic*, Vol. 19, No. 3, pp. 351–393, September 2013.

[38] Koellner, Peter, "Truth in Mathematics: The Question of Pluralism", *New Waves in Philosophy of Mathematics*, pp. 80–116, O. Bueno & Ø. Linnebo (eds.), Palgrave Macmillan, 2009.

[39] Kornblith, Hilary, "Epistemic Normativity", *Synthese*, Vol. 94, pp. 357–376, 1993.

[40] Kouri Kissel, Teresa, *Logical Instrumentalism* (Doctoral thesis, Ohio State University), 2016.

[41] Kouri Kissel, Teresa, "Logical Pluralism from a Pragmatic Perspective", *Australian Journal of Philosophy*, November 2017.

[42] Kripke, Saul, "Outline of a Theory of Truth" *The Journal of Philosophy* Vol. 72, No. 19, pp. 690–716, Journal of Philosophy, Inc., November 1975.

[43] Lau, Tszyuen and Wang, Yanjing, "Knowing Your Ability", *The Philosophical Forum*, Vol. 47, Issue 3-4, pp. 415–423, Fall-Winter 2016.

[44] Leitgeb, Hannes, "What theories of truth should be like (but cannot be)", *Philosophy Compass*, 2/2, pp. 276–290, 2007.

[45] Lynch, Michael P., *Truth as One and Many*, Oxford University Press, 2009.

[46] MacFarlane, John, *What Does It Mean That Logic Is Formal?* (PhD thesis, University of Pittsburgh), 2000.

[47] MacFarlane, John, "In What Sense (If Any) Is Logic Normative for Thought?", unpublished manuscript.

[48] MacFarlane, John, "Is Logic a Normative Discipline?", unpublished manuscript.

[49] Misak, C. J., *Truth and the End of Inquiry: A Peircean Account of Truth*, Oxford University Press, 2004.

[50] Morton, Adam, "Denying the Doctrine and Changing the Subject", *The Journal of Philosophy*, Vol. 70, No. 15, pp. 503–510, 1973.

[51] Pedersen, Nikolaj Jang Lee Linding, "Pluralism × 3: Truth, Logic, Metaphysics" *Erkenntnis*, Vol. 79, Supplement 2, pp. 259–277, March 2014.

[52] Priest, Graham, "The Logic of Paradox", *Journal of Philosophical Logic*, Vol. 8, pp. 219–241, 1979.

[53] Priest, Graham, "Hyper-contradictions", *Logic et Analyse*, Vol. 27, pp. 237–243, 1984.

[54] Priest, Graham, "The Logic of Paradox Revisited", *Journal of Philosophical Logic*, Vol. 13, No. 2, pp. 153–179, 1984.

[55] Priest, Graham, *In Contradiction: A Study of the Transconsistent*, Expanded Edition, Oxford University Press, 2006.

[56] Priest, Graham, *Doubt Truth to Be a Liar*, Oxford University Press, 2006.

[57] Priest, Graham, "Logical Disputes and the a Priori", *Logique & Analyse*, Vol. 59, No. 236, 2015.

[58] Quine, Willard Van Orman, *Philosophy of Logic*, Harvard University Press, June 1986.

[59] Rahman, Shahid and Redmond, Juan, "Hugh MacColl and the Birth of Logical Pluralism", In Dov M. Gabbay and John Woods (eds.), *Handbook of the History of Logic, Volume 4: British Logic in the Nineteenth Century*, pp. 533–604, 2008.

[60] Ramsay, Frank, "Truth and Probability", *The Foundations of Mathematics and Other Logical Essays*, pp. 156–198, R. B. Braithwaite (ed.), Kegan, Paul, Trench, Trubner & Co., 1931.

[61] Read, Stephen, "Anti-Exceptionalism about Logic", *Australasian Journal of Logic*, Vol. 16, No. 7, 2019.

[62] Roberts, Craige, "Information Structure in Discourse: Towards an Integrated Formal Theory of Pragmatics", *Semantics & Pragmatics* Vol. 5, Art. 6 pp. 1–69, 2012.

[63] Russell, Bertrand, "What the Tortoise Said to Achilles", *The Principles of Mathematics*, Cambridge University Press, 1903.

[64] Russell, Gillian, "Logic Isn't Normative", *Inquiry*, September 2017.

[65] Russell, Gillian, "Deviance And Vice: Strength As A Theoretical Virtue In The Epistemology Of Logic", *Philosophy and Phenomenological Research*, Vol. 99, No. 3, pp. 548-563, 2019.

[66] Ryle, Gilbert, "Knowing How and Knowing That", *Collected Papers, Volume 2: Collected Essays 1929–1968*, pp. 222–236, Routledge, 2009.

[67] Scharp, Kevin, *Replacing Truth*, Oxford University Press, 2013.

[68] Schlenker, Philippe, "Super Liars", *The Review of Symbolic Logic* Vol. 3 No. 3, pp. 374–414, 2010.

[69] Shapiro, Stewart, *Varieties of Logic*, Oxford University Press, 2014.

[70] Shapiro, Stewart and Kouri Kissel, Teresa, "Logical Pluralism and Normativity", *Inquiry*, August 2017.

[71] Sosa, Ernest, *Judgment and Agency*, Oxford University Press, 2015.

[72] Stalnaker, Robert, *Context*, Oxford University Press, 2014.

[73] Stanley, Jason, & Williamson, Timothy, "Knowing How", *Journal of Philosophy*, Vol. 98, No. 8, 2001.

[74] Steinberger, Florian, "Three Ways in which Logic Might Be Normative" *The Journal of Philosophy*, Vol. 116, No. 1, pp. 5–31, 2019.

[75] Stich, Stephen, The Fragmentation of Reason, MIT Press, 1990.

[76] Stich, Stephen & Nisbett, Richard E., "Justification and the Psychology of Human Reasoning", *Philosophy of Science*, Vol. 47, pp. 188–202, 1980.

[77] Tarski, Alfred, *Introduction to Logic and the Methodology of the Deductive Sciences*, fourth edition, Jan Tarski (ed.), Oxford University Press, 1994.

[78] Tarski, Alfred, "On the Concept of Following Logically", *History and Philosophy of Logic*, Vol. 23, pp. 155–196, M. Stroinska & D. Hitchcock (trans.), Oxford University Press, 2002.

[79] Thagard, Paul, *Computational Philosophy of Science*, MIT Press, 1988.

[80] Thagard, Paul, "Why Cognitive Science Needs Philosophy and Vice Versa", *Topics in Cognitive Science*, Vol. 1, No. 2, pp. 237–254, April 2009.

[81] Troelstra, Anne, *History of Constructivism in the Twentieth Century*, ITLI Prepublication Series for Mathematical Logic and Foundations, 1991.

[82] van Benthem, Johan, "Logic and Philosophy, a Sea of Stories", *Tsinghua Studies in Western Philosophy*, Vol. 1, No. 1, pp. 124–153, 2015.

[83] Varzi, Achille C., "On Logical Relativity", *Philosophical Issues* Vol. 10, pp. 197–219, 2002.

[84] Visser, Albert, "Four Valued Semantics and the Liar", *Journal of Philosophical Logic*, Vol. 13 No. 2, pp. 181–212, 1984.

[85] Visser, Albert, "Interpretations over Heyting's Arithmetic", *Logic at Work, Studies in Fuzziness and Soft Computing*, pp. 255–284, E. Orłowska (ed.), 1998.

[86] Wang, Yanjing, "A Logic of Goal-Directed Knowing How", *Synthese*, Vol. 195, pp. 4419–4439, 2018.

[87] Weber, Zach, "Transfinite cardinals in paraconsistent set theory", *Review of Symbolic Logic*, Vol. 5, No. 2, pp. 269–293, 2012.

[88] Williamson, Jon, *In Defense of Objective Bayesianism*, Oxford University Press, 2010.

[89] Williamson, Timothy, *Knowledge and Its Limits*, Oxford University Press, 2002.

[90] Williamson, Timothy, "Philosophical 'Intuitions' and Scepticism about Judgement", *Dialectica*, Vol. 58, No. 1, pp. 109–153, 2004.

[91] Williamson, Timothy, *Modal logic as metaphysics*, Oxford University Press, 2013.

[92] Williamson, Timothy, "Logic, Metalogic and Neutrality", *Erkenntnis* Vol. 79, Supplement 2, pp. 211–231, March 2014.

[93] Williamson, Timothy, "Semantic Paradoxes and Abductive Methodology", *Reflections on the Liar*, Bradley Armour-Garb (ed.), pp. 325–346, Oxford University Press, 2017.

[94] Williamson, Timothy, "Alternative Logics and Applied Mathematics", *Philosophical Issues*, Vol. 28, pp. 399–424 5 March 2018.

[95] Woods, Jack, "Against Reflective Equilibrium for Logical Theorizing", *The Australasian Journal of Logic*, Vol. 16, No. 7, 2019.

[96] Yaqūb, Aladdin, *The Liar Speaks the Truth*, Oxford University Press, 1993.

Samenvatting

Een Pragmatische Verdediging van het Logisch Pluralisme

In dit proefschrift wordt logica opgevat als een formele presentatie van een leidraad bij een of andere redeneerpraktijk, een leidraad die zelf de uitkomst is van epistemische normen en hun gevolgen. Dit pragmatisch gezichtspunt biedt een basis voor logisch pluralisme, de opvatting dat er meer dan één goede, juiste of correcte logica is. Er kan immers meer dan één goede redeneerpraktijk bestaan, en meer dan één wijze om die praktijk te concipiëren. Deze vorm van logisch pluralisme noemen we alomvattend logisch pluralisme.

In het eerste hoofdstuk van het proefschrift wordt geschetst wat er bedoeld wordt met 'formeel' in 'formele presentatie', en wat met 'epistemische normen'. Bovendien wordt in het eerste hoofdstuk aangetoond dat het alomvattend logisch pluralisme bestand is tegen de gebruikelijke kritiek op logisch pluralisme, die luidt dat het logisch pluralisme ten onder gaat aan interne contradicties of dat het vervalt in logisch monisme. Ten slotte worden enkele andere benaderingen van logisch pluralisme besproken en vergeleken met de onze, waaronder die van J. C. Beall en Greg Restall, die van Nikolaj Pedersen en Michael P. Lynch, en die van Stewart Shapiro en Teresa Kouri Kissel.

In het tweede hoofdstuk van het proefschrift wordt de opvatting verdedigd dat het doel van logische revisie is een reflectief evenwicht te bereiken tussen de epistemische normen die de praktijk bepalen die door de logica in kwestie geformaliseerd wordt aan de ene kant en de normen van de formalisering zelf aan de andere. Dit standpunt wordt verdedigd tegen enkele invloedrijke bezwaren, en gecontrasteerd met een aantal recente opvattingen over logische revisie die varen onder de vlag van anti-exceptionalisme (waaronder de opvatting van Timothy Williamson, die van Ole Thomassen Hjortland en die van Graham Priest). Het logisch anti-exceptionalisme blijkt onverenigbaar te zijn met de pragmatische opvatting van logica die in het vorige hoofdstuk is geschetst.

Het derde hoofdstuk van het proefschrift schetst een aantal nauw verwante oplossingen voor de semantische paradoxen — dit geheel in lijn met de bovenstaande opvatting over logische revisie: de beoefenaar identificeert de epistemische normen die hij wenst te behouden en past ofwel zijn praktijk van redeneren ofwel de formalisering daarvan dienovereenkomstig aan. De benadering is zo algemeen dat het derde hoofdstuk eigenlijk een soort een case study is. Meer specifiek wordt in navolging van Roy T. Cook het bestaan van zogenaamde 'revanche' paradoxen gezien als argument voor de stelling dat het begrip waarheidswaarde onbepaald uitbreidbaar (in de zin van M. Dummett) is. Zo'n uitbreiding is nodig wanneer we de predikaten waarmee zulke paradoxen geformuleerd kunnen worden, zoals *'waar en paradoxaal'*, in onze taal opnemen. Er wordt bewezen dat er Kripkeaanse dekpunten bestaan voor een ordinaalwaardige hiërarchie van zulke predikaten —dit als een formele uitwerking van het verschijnsel revanche paradox.

Het vierde hoofdstuk van het proefschrift beschrijft hoe beoefenaars van verschillende logica's elkaars bewijzen kunnen begrijpen en van elkaars bewijzen kunnen leren, voorzover er tenminste geen linguistische beperkingen zijn die een formele vertaling van de ene logica in de andere in de weg staan. Er wordt een opvatting van het begrip 'interpretatie' verdedigd die gebaseerd is op een praktijkgericht *principe van welwillendheid*: We interpreteren anderen het best als we veronderstellen dat ze epistemische normen volgen die maximaal epistemisch nut hebben bij wat ze in onze ogen willen bewerkstelligen. Dit principe speelt een belangrijke rol bij een pragmatiek voor communicatie voor beoefenaars van verschillende logica's. Deze pragmatiek wordt vorm gegeven door het vermogen van de wiskundige om de relevante gevolgtrekkingen te maken, te karakteriseren binnen een variant van Craige Roberts' *Questions under Discussion* raamwerk. Het resulterende raamwerk wordt toegepast op een aantal (imaginaire) voorbeelden van discussies tussen beoefenaars van verschillende logica's, zoals die tussen een intuitionist en een klassiek geörienteerde wiskundige.

Summary

A Pragmatic Defense of Logical Pluralism

In this thesis, logic is conceived as a formal presentation of a guide to undertaking a rational practice, a guide which is itself constituted by epistemic norms and their consequences. This pragmatic conception of logic is a basis for logical pluralism, the view that there is more than one good, right, or correct logic. After all there may be more than one good practice, and more than one way to conceive the practice. This conception of logical pluralism we refer to as thoroughgoing logical pluralism.

The first chapter of the thesis outlines what we mean by "formal" in "formal presentation", and what we mean by "epistemic norms". Furthermore, the first chapter shows how thoroughgoing logical pluralism holds up against the usual challenges to logical pluralism, namely that logical pluralism is self-defeating and that logical pluralism collapses into logical monism. Finally, some other approaches to logical pluralism are accounted for and contrasted with the present account, including those of J. C. Beall and Greg Restall, Nikolaj Pedersen and Michael P. Lynch, and Stewart Shapiro and Teresa Kouri Kissel.

The second chapter of the thesis defends the view that the goal of logical revision is merely reflective equilibrium with respect to the norms that constitute the practice that the logic is formalising, along with the norms of the formalisation itself. This stance is defended against some influential objections, and contrasted with a number of recent accounts under the umbrella of logical anti-exceptionalism (including Timothy Williamson's, Ole Thomassen Hjortland's, and Graham Priest's). The approach of logical anti-exceptionalism in general is shown to be incompatible with the pragmatic conception of logic outlined in the previous chapter.

The third chapter of the thesis outlines a family of solutions to the semantic paradoxes in line with the preceding account of logical revision: the practitioner identifies the norms that they wish to keep, and revise their practice or formalisation accordingly. The approach is sufficiently general that the third chapter is really more of a case study. In particular, in line with what Roy T. Cook has observed, revenge liar paradoxes can be reconceived as demonstrating that the concept of truth value is indefinitely extensible (in Michael Dummett's sense) when our truth-talk includes predicates (such as those for "true and false and 'true and false'") that allow us to formulate these paradoxes. Kripkean fixed point extensions are shown to exist for an ordinal-valued hierarchy of truth-and-paradoxicality predicates to support a formal articulation of the phenomenon of revenge liar paradoxes, insofar as we may wish to articulate that phenomenon.

The fourth chapter of the thesis shows an account of how practitioners of different logics may understand and learn from each-other's proofs, given that a formal translation of one logic inside another that is sufficiently informative to present this account is allowed within linguistic constraints. A notion of interpretation is defended based on a practice-oriented principle of charity: that we make the best sense of others when we suppose they are following epistemic norms with maximal epistemic utility with respect to our possible interpretations of what their instrumental (goal-directed) desires could be. This notion of interpretative charity is employed in a pragmatics of communication for practitioners of different logics. This pragmatics is managed with a variation of Craige Roberts' Questions Under Discussion framework, characterising the potential ability of the formal mathematician to make the relevant inferences. This framework is applied to a few examples of conversations between practitioners, such as between a hypothetical intuitionistic and classical mathematician.

Titles in the ILLC Dissertation Series:

ILLC DS-2016-01: **Ivano A. Ciardelli**
Questions in Logic

ILLC DS-2016-02: **Zoé Christoff**
Dynamic Logics of Networks: Information Flow and the Spread of Opinion

ILLC DS-2016-03: **Fleur Leonie Bouwer**
What do we need to hear a beat? The influence of attention, musical abilities, and accents on the perception of metrical rhythm

ILLC DS-2016-04: **Johannes Marti**
Interpreting Linguistic Behavior with Possible World Models

ILLC DS-2016-05: **Phong Lê**
Learning Vector Representations for Sentences - The Recursive Deep Learning Approach

ILLC DS-2016-06: **Gideon Maillette de Buy Wenniger**
Aligning the Foundations of Hierarchical Statistical Machine Translation

ILLC DS-2016-07: **Andreas van Cranenburgh**
Rich Statistical Parsing and Literary Language

ILLC DS-2016-08: **Florian Speelman**
Position-based Quantum Cryptography and Catalytic Computation

ILLC DS-2016-09: **Teresa Piovesan**
Quantum entanglement: insights via graph parameters and conic optimization

ILLC DS-2016-10: **Paula Henk**
Nonstandard Provability for Peano Arithmetic. A Modal Perspective

ILLC DS-2017-01: **Paolo Galeazzi**
Play Without Regret

ILLC DS-2017-02: **Riccardo Pinosio**
The Logic of Kant's Temporal Continuum

ILLC DS-2017-03: **Matthijs Westera**
Exhaustivity and intonation: a unified theory

ILLC DS-2017-04: **Giovanni Cinà**
Categories for the working modal logician

ILLC DS-2017-05: **Shane Noah Steinert-Threlkeld**
Communication and Computation: New Questions About Compositionality

ILLC DS-2017-06: **Peter Hawke**
The Problem of Epistemic Relevance

ILLC DS-2017-07: **Aybüke Özgün**
Evidence in Epistemic Logic: A Topological Perspective

ILLC DS-2017-08: **Raquel Garrido Alhama**
Computational Modelling of Artificial Language Learning: Retention, Recognition & Recurrence

ILLC DS-2017-09: **Miloš Stanojević**
Permutation Forests for Modeling Word Order in Machine Translation

ILLC DS-2018-01: **Berit Janssen**
Retained or Lost in Transmission? Analyzing and Predicting Stability in Dutch Folk Songs

ILLC DS-2018-02: **Hugo Huurdeman**
Supporting the Complex Dynamics of the Information Seeking Process

ILLC DS-2018-03: **Corina Koolen**
Reading beyond the female: The relationship between perception of author gender and literary quality

ILLC DS-2018-04: **Jelle Bruineberg**
Anticipating Affordances: Intentionality in self-organizing brain-body-environment systems

ILLC DS-2018-05: **Joachim Daiber**
Typologically Robust Statistical Machine Translation: Understanding and Exploiting Differences and Similarities Between Languages in Machine Translation

ILLC DS-2018-06: **Thomas Brochhagen**
Signaling under Uncertainty

ILLC DS-2018-07: **Julian Schlöder**
Assertion and Rejection

ILLC DS-2018-08: **Srinivasan Arunachalam**
Quantum Algorithms and Learning Theory

ILLC DS-2018-09: **Hugo de Holanda Cunha Nobrega**
Games for functions: Baire classes, Weihrauch degrees, transfinite computations, and ranks

ILLC DS-2018-10: **Chenwei Shi**
Reason to Believe

ILLC DS-2018-11: **Malvin Gattinger**
New Directions in Model Checking Dynamic Epistemic Logic

ILLC DS-2018-12: **Julia Ilin**
Filtration Revisited: Lattices of Stable Non-Classical Logics

ILLC DS-2018-13: **Jeroen Zuiddam**
Algebraic complexity, asymptotic spectra and entanglement polytopes

ILLC DS-2019-01: **Carlos Vaquero**
What Makes A Performer Unique? Idiosyncrasies and commonalities in expressive music performance

ILLC DS-2019-02: **Jort Bergfeld**
Quantum logics for expressing and proving the correctness of quantum programs

ILLC DS-2019-03: **András Gilyén**
Quantum Singular Value Transformation & Its Algorithmic Applications

ILLC DS-2019-04: **Lorenzo Galeotti**
The theory of the generalised real numbers and other topics in logic

ILLC DS-2019-05: **Nadine Theiler**
Taking a unified perspective: Resolutions and highlighting in the semantics of attitudes and particles

ILLC DS-2019-06: **Peter T.S. van der Gulik**
Considerations in Evolutionary Biochemistry

ILLC DS-2019-07: **Frederik Möllerström Lauridsen**
Cuts and Completions: Algebraic aspects of structural proof theory

ILLC DS-2020-01: **Mostafa Dehghani**
Learning with Imperfect Supervision for Language Understanding

ILLC DS-2020-02: **Koen Groenland**
Quantum protocols for few-qubit devices

ILLC DS-2020-03: **Jouke Witteveen**
Parameterized Analysis of Complexity

ILLC DS-2020-04: **Joran van Apeldoorn**
A Quantum View on Convex Optimization

ILLC DS-2020-05: **Tom Bannink**
Quantum and stochastic processes

ILLC DS-2020-06: **Dieuwke Hupkes**
Hierarchy and interpretability in neural models of language processing

ILLC DS-2020-07: **Ana Lucia Vargas Sandoval**
On the Path to the Truth: Logical & Computational Aspects of Learning

ILLC DS-2020-08: **Philip Schulz**
Latent Variable Models for Machine Translation and How to Learn Them

ILLC DS-2020-09: **Jasmijn Bastings**
A Tale of Two Sequences: Interpretable and Linguistically-Informed Deep Learning for Natural Language Processing

ILLC DS-2020-10: **Arnold Kochari**
Perceiving and communicating magnitudes: Behavioral and electrophysiological studies

ILLC DS-2020-11: **Marco Del Tredici**
Linguistic Variation in Online Communities: A Computational Perspective

ILLC DS-2020-12: **Bastiaan van der Weij**
Experienced listeners: Modeling the influence of long-term musical exposure on rhythm perception

ILLC DS-2020-13: **Thom van Gessel**
Questions in Context

ILLC DS-2020-14: **Gianluca Grilletti**
Questions & Quantification: A study of first order inquisitive logic

ILLC DS-2020-15: **Tom Schoonen**
Tales of Similarity and Imagination. A modest epistemology of possibility

ILLC DS-2020-16: **Ilaria Canavotto**
Where Responsibility Takes You: Logics of Agency, Counterfactuals and Norms