2-2023

Three Essays on Substructural Approaches to Semantic Paradoxes

Brian C. Porter
The Graduate Center, City University of New York

Three Essays on Substructural Approaches to Semantic Paradoxes

by

Brian Cross Porter

A dissertation submitted to the Graduate Faculty in Philosophy in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2022

This manuscript has been read and accepted by the Graduate Faculty in Philosophy in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Professor Nickolas Pappas

Date — Chair of Examining Committee

Professor Nickolas Pappas

Date — Executive Officer

Professor Graham Priest

Professor Melvin Fitting

Professor Hartry Field

Professor Eduardo Barrio

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

Three Essays on Substructural Approaches to Semantic Paradoxes

by

Brian Cross Porter

Adviser: Professor Graham Priest

This dissertation consists of three papers:

Supervaluations and the Strict-Tolerant Hierarchy

In a recent paper, Barrio, Pailos and Szmuc (BPS) show that there are logics that have exactly the validities of classical logic up to arbitrarily high levels of inference. They suggest that a logic therefore must be identified by its valid inferences at every inferential level. However, Scambler shows that there are logics with all the validities of classical logic at every inferential level, but with no antivalidities at any inferential level. Scambler concludes that in order to identify a logic, we at least need to look at the validities and the antivalidities of every inferential level. In this paper, I argue that this is still not enough to identify a logic. I apply BPS's techniques in a super/sub-valuationist setting to construct a logic that has exactly the validities and antivalidities of classical logic at every inferential level. I argue that the resulting logic is nevertheless distinct from classical logic.

Infinite-Premise Paradoxes and Contraction

In recent years, there have been several defenses of noncontractive solutions to semantic paradoxes. These solutions propose that we avoid paradox by dropping the structural rule

of *Contraction*, and allow that a premise taken twice have different consequences than the premise taken only once. This repeats for every n: a premise taken *n+1* times has different consequences than a premise taken only n times. This solution runs into trouble when we consider contexts that allow for infinitely many premises, since $\aleph_0 + 1 = \aleph_0$. In this paper, I introduce infinitary versions of several semantic paradoxes, which involve infinitely many premises and therefore cannot be solved by noncontractive approaches. I argue that non-contractive solutions cannot avoid this problem, and that noncontractive solutions therefore fail to provide a uniform solution to semantic paradoxes.

An Infinite Hierarchy of Validity Curry Paradoxes

A version of Curry's paradox involving a validity predicate (hereafter *the validity curry paradox*) has in recent years been presented as a particular problem for non-classical theories of truth that don't want to go substructural. In this paper, I introduce an infinite hierarchy of validity curry paradoxes that apply to *metainferential* validity. None of these paradoxes can be solved by giving up the usual Cut or Contraction. I argue that any attempt to block every paradox in the hierarchy at the same step faces serious problems. In particular, the substructuralist must reject Cut or Contraction at every metainferential level, which has worse results than rejecting either rule at the first level alone. I propose a solution that keeps all of the usual structural rules, but rejects different validity rules at different levels. I argue that this is nevertheless a uniform solution.

Acknowledgments

The papers presented here only exist because I have had the good fortune to receive the support of many people.

I owe an immense debt to Graham Priest, my supervisor, who has been consistently generous with his time and advice for as long I've known him. Virtually every idea that I present this thesis—and a great many more ideas than that—can be traced back to a question I posed to Graham. A question to which he inevitably responded promptly, with deep insights and suggestions to develop that idea. Without his support and encouragement, I could never have written this thesis. I thank him for all of his help and guidance. It has been invaluable.

I would also like to thank Eduardo Barrio, Hartry Field, and Melvin Fitting. Their feedback has been immensely helpful, and has made this thesis possible. I appreciate their kindness, patience and generosity. The work presented in this thesis would not exist if I had not had the opportunity to learn from them. I owe a wide range of intellectual debts to each of them. They have taught me how logic and philosophy ought to be done. I thank them for their support, and for the examples they have set.

I would not have been able to write this thesis without the patience and understanding of my partner, Alondra Clabaguera Fulgencio. Her care and support have sustained me throughout this process. I could not have done this without her.

I would like to thank Michael Devitt, my coauthor of two papers, who has taught me more about passion and work ethic than anyone else.

I would like to thank Jen Marušić, who taught me to develop a paper, and who taught me how to follow an idea wherever it takes you. I would not be here without her kindness and generosity.

Finally, I want to thank Jim Cargile, Daniel Devereux, and Trenton Merricks. Together, the three of them set me on the path that led to this thesis, and to academic philosophy as a career. Their courses convinced me that I wanted to study philosophy, and their encouragement gave me hope that I could do it for a living. They opened up whole new worlds to me, and first taught me how to think through issues and construct arguments with care and rigor.

Permissions

The present dissertation incorporates material that has previously appeared in print. I am grateful to Springer Nature for permission to include this material.

Chapter 2, "Supervaluations and the Strict-Tolerant Hierarchy", is a revised version of the paper:

Porter, B. (2021). Supervaluations and the Strict-Tolerant Hierarchy. *Journal of Philosophical Logic* https://doi.org/10.1007/s10992-021-09624-6.

Contents

Chapter 1

Introduction

The papers presented here examine substructural approaches to semantic paradoxes. Substructural approaches to semantic paradoxes advocate dropping one of the *structural rules*, usually either Cut or Contraction, in order to solve paradoxes of truth and validity. I will use a sequent calculus presentation throughout, as a sequent calculus presentation will make it easier to see the structural rules used in derivations, and will also make it easier to see the role of metainferences in our later investigation of metainferential paradoxes. We are interested in validity, and the sequent calculus provides a convenient way to prove which inferences are valid.

The classical sequent calculus LK involves sequents of the form $\Gamma \Rightarrow \Delta$. Both Γ and Δ are understood to be *multisets*. Therefore $\Gamma, A \Rightarrow \Delta$ is equivalent to $A, \Gamma \Rightarrow \Delta$, but $\Gamma, A, A \Rightarrow \Delta$ is not equivalent to $\Gamma, A \Rightarrow \Delta$.

LK has two sorts of rule: structural rules, and operational rules.

The structural rules are Identity, Weakening, Cut, and Contraction:

$$\frac{}{A \Rightarrow A}\text{ Identity}$$

$$\frac{\Gamma \Rightarrow \Delta}{\Gamma, A \Rightarrow \Delta}\text{ Weakening (left)}$$

$$\frac{\Gamma \Rightarrow \Delta}{\Gamma \Rightarrow B, \Delta}\text{ Weakening (right)}$$

$$\frac{\Gamma, A, A \Rightarrow \Delta}{\Gamma, A \Rightarrow \Delta}\text{ Contraction (left)}$$

$$\frac{\Gamma \Rightarrow B, B, \Delta}{\Gamma \Rightarrow B, \Delta}\text{ Contraction (right)}$$

$$\frac{\Gamma, A \Rightarrow \Delta \qquad \Gamma' \Rightarrow A, \Delta'}{\Gamma, \Gamma' \Rightarrow \Delta, \Delta'}\text{ Cut}$$

Any logic that drops one of these structural rules is a *substructural* logic. Logics that drop Identity are called “irreflexive”; logics that drop Weakening are “nonmonotonic”; logics that drop Contraction are called "noncontractive"; and logics that drop Cut are called "nontransitive". Noncontractive and nontransitive logics have been much more popular as solutions to semantic paradoxes than irreflexive logics, and so I will primarily focus on those two approaches. Nonmonotonic logics have been defended,[1] and some of the noncontractive logics put forward as solutions to semantic paradoxes have been also been nonmonotonic. But since Identity and Weakening are not directly involved in the derivations of every paradox—in particular, they appear to play no role in the validity curry paradox—irreflexive and nonmonotonic approaches do not automatically solve all of the semantic paradoxes that concern us.

In addition to structural rules, *LK* also has *operational* rules. These are rules that introduce logical operators like ¬ and ∨. Each operator has two rules: one left-side rule, and one right-side rule. These are:

$$\frac{\Gamma \Rightarrow A, \Delta}{\Gamma \neg A \Rightarrow \Delta}\ \neg\mathrm{L}$$

$$\frac{\Gamma, A \Rightarrow \Delta}{\Gamma \Rightarrow \neg A, \Delta}\ \neg\mathrm{R}$$

$$\frac{\Gamma, A \Rightarrow \Delta \qquad \Gamma', B \Rightarrow \Delta'}{\Gamma, \Gamma', A \vee B \Rightarrow \Delta, \Delta'}\ \vee\mathrm{L}$$

$$\frac{\Gamma \Rightarrow A, B, \Delta}{\Gamma \Rightarrow A \vee B, \Delta}\ \vee\mathrm{R}$$

$$\frac{\Gamma A, B \Rightarrow \Delta}{\Gamma A \wedge B \Rightarrow \Delta}\ \wedge\mathrm{L}$$

$$\frac{\Gamma \Rightarrow A, \Delta \qquad \Gamma' \Rightarrow B, \Delta'}{\Gamma, \Gamma' \Rightarrow A \wedge B, \Delta, \Delta'}\ \wedge\mathrm{R}$$

$$\frac{\Gamma \Rightarrow A, \Delta \qquad \Gamma', B \Rightarrow \Delta'}{\Gamma, \Gamma', A \rightarrow B \Rightarrow \Delta, \Delta'}\ \rightarrow\mathrm{L}$$

[1] See for example [8].

$$\frac{\Gamma, A \Rightarrow B, \Delta}{\Gamma \Rightarrow A \to B, \Delta} \to\text{R}$$

With these structural and operational rules, we can derive a sequent $\Gamma \Rightarrow \Delta$ in LK iff $\Gamma \vdash \Delta$ is a valid inference in classical logic. For example, we can derive classically valid inferences like modus ponens:

$$\frac{\dfrac{}{A \Rightarrow A}\text{Identity} \qquad \dfrac{}{B \Rightarrow B}\text{Identity}}{A, A \to B \Rightarrow B} \to\text{L}$$

the Law of Excluded Middle:

$$\dfrac{\dfrac{\dfrac{}{A \Rightarrow A}\text{Identity}}{\Rightarrow A, \neg A}\neg\text{R}}{\Rightarrow A \vee \neg A}\vee\text{R}$$

and Explosion:

$$\dfrac{\dfrac{\dfrac{\dfrac{}{A \Rightarrow A}\text{Identity}}{A, \neg A \Rightarrow}\neg\text{L}}{A \wedge \neg A \Rightarrow}\wedge\text{L}}{A \wedge \neg A \Rightarrow B}\text{Weakening (right)}$$

As a result, non-classical approaches that aim to solve semantic paradoxes by dropping LEM or Explosion will have to restrict or reject some of the rules of LK. However, different paradoxes involve different connectives: the curry paradox involves a conditional, the liar paradox involves negation, and the validity curry predicate involves neither. This has led defenders of the substructural approach to argue that other non-classical approaches fail to provide a uniform solution to semantic paradoxes.

In our attempts to solve semantic paradoxes, much importance is placed on the notion of *uniformity*. As Priest phrases it in [7]: "same kind of paradox, same kind of solution."

One of the appeals of substructural approaches to semantic paradoxes is that they appear to provide undeniably uniform solutions: every paradoxical argument is blocked at the exact same step. Thus regardless of which paradoxes are "of the same kind", substructural approaches are guaranteed to provide the "same kind" of solution to those paradoxes.

Defenders of substructural approaches argue that other non-classical approaches do not provide a uniform solution, because they must block negation rules for the liar, conditional rules for the curry, and so on. See, for example, Ripley [13]:

> "The real problem with [nonsubstructural approaches]... is this: the nonsubstructuralist deals with the paradoxes piecemeal, missing the general features that are allowing them to arise in the first place. But paradox runs deeper than any particular vocabulary. Tinkering with negation or conditional rules might prevent paradoxes involving negations and conditionals from arising, but it doesn't come to grips with the general phenomenon."

The charge again non-substructural approaches, then, is that they do not solve the general phenomenon underlying the semantic paradoxes. They provide specific solutions to specific paradoxes, but fail to provide any solution that solves the paradoxes in full generality.

However, I argue that it is *substructural* approaches that do not "come to grips with the general phenomenon", because substructural logics simply ignore many of the contexts in which we want to use logics, and in which paradoxes can arise.

We use logic, and we use logics, for many purposes. And for many purposes, substructural logics can be very useful.

For example, Ripley [12] gives an bilateralist account of the nontransitive ST consequence relation, such that $\Gamma \vdash \Delta$ is understood to mean that it is incoherent to assert all of the γs while denying all of the δs. On this understanding, Cut fails precisely for sentences that cannot be coherently asserted or denied. This may be a useful consequence relation that can serve valuable purposes. In [4], Kripke suggests that paradoxical and ungrounded sentences like the liar sentence fail to express a proposition. One may take assertion and denial to essentially involve assertion or denial of *propositions*, and thus take the liar sentence to be neither assertible nor deniable. In that case, we might think that ST gives the right consequence relation *for determining which sets of assertions and denials are impermissible.* And we furthermore might take this to be valuable to give an inferentialist account of the meanings of the logical connectives. In this way, the nontransitive consequence relation might be a very useful tool.

Similarly, Mares and Paoli [5] describe one sense in which we might say that A "follows from" Γ. Namely, that A follows from Γ if "given the rules of the logic at issue, we can extract the information that A from the combined information provided by the sentences in Γ" [5]. They argue that the structural rule of Contraction fails for this notion of "follows from", because given the rules of a particular logic, we sometimes need more than one application of a particular premise in the proof of a theorem. There may be a proof of B that uses A twice, and there may be no proof of B that uses A only once. In such a case, structural Contraction fails: given the rules of the logic at issue, two applications of A suffice to prove B, but one does not. This notion of consequence avoid semantic paradoxes because, for example, we need two applications of the curry sentence in order to derive an arbitrary sentence; there is no proof of an arbitrary sentence that only requires one application of a curry sentence. I think this notion of "follows from" is a perfectly legitimate notion, and that a noncontractive notion of consequence may therefore be very useful to investigate these sorts of features of a particular logic.

Nontransitive and noncontractive consequence relations are well-suited to serve these purposes. However, there are many other purposes for which we want a logic, and substructural logics simply cannot serve all of these purposes. One of the most important uses of a logic is to discover the consequences of a collection of premises. Given a collection of axioms, or assumptions, or beliefs, we want to know what those axioms/assumptions/beliefs *commit* us to. This is an important part of how we use logic informally, in our daily lives, but it is also an important use of formal logics in mathematics and other areas. We want to know whether or not Goldbach's Conjecture is a consequence of the axioms of ZFC, for example.

One important role of a logic is to provide a tool that can be used to discover these sorts of consequences.

For a logic to serve this purpose, the consequence relation of the logic must form a *closure operator*. A closure operator C(X) is an operator on sets (or multisets) such that:

1. $X \subseteq C(X)$
2. if $X \subseteq Y$ then $C(X) \subseteq C(Y)$
3. $C(X) = C(C(X)$

If a consequence relation $\vdash$ corresponds to a closure operator $C(X)$, such that $\Gamma \vdash A$ iff $A \in C(\Gamma)$, then it must obey the usual structural rules.

(1) gives us identity: $A \vdash A$

(2) gives us monotonicity: $\Gamma \vdash A$ implies $\Gamma, B \vdash A$

(3) gives us "Cautious Cut": $\Gamma \vdash A$ and $\Gamma, A \vdash B$ implies $\Gamma \vdash B$.

Together, (1), (2), and (3) make both Cut and Contraction derivable rules in our logic. As such, no nontransitive or noncontractive logics can correspond to a closure operator. This has several problems, one of which is that we cannot use such a logic to discover the consequences of a collection of axioms.

Normally we can take e.g. the axioms of ZFC, and know that there is a singular set of theorems provable from those axioms. This is the *closure* of the axioms. To quote Beall: "Give to logic your theory Γ, and then sit back: logic 'freely' or 'automatically' expands your theory to $C(\Gamma)$, which contains all of Γ's (singleton) consequences" [2].

This power of a logic is dependent on it corresponding to a closure operator. Without a closure operator, we cannot use a logic to present the entirety of a theory using just a set of non-logical axioms. Therefore, whatever other benefits they may provide, noncontractive and nontransitive logics do not have this power.

The traditional view is that a logic is a consequence relation, where a consequence relation is a set of valid inferences. This can be at least partially traced back to Tarski. However, the notion of consequence relation that Tarski uses is that of a single-conclusion consequence relation that is reflexive, transitive, and monotonic; see for example [15]. This "Tarskian" notion of consequence relation simply does not apply to substructural logics. As Barrio et al. discuss in [1] in the context of nontransitive logics, and as Cintula and Paoli discuss in [3] in the context of noncontractive logics, substructural consequence relations are not Tarskian consequence relations. If we want to discuss substructural logics, we must "move beyond the Tarksian paradigm", as Barrio et al. put it in [1].

Moving beyond the Tarskian paradigm has many consequences. One consequence, which I discuss in the first paper presented here, is that we can no longer identify a logic just by its set of valid inferences. Barrio et al., Scambler, and Fitting paint a clear picture: given a logic with a Tarskian consequence relation like that of classical logic, K3, LP, or FDE, and there is a substructural logic with the same valid inferences that nevertheless differs in other ways.

Another consequence is that not every logic we may consider can be used for all of the purposes for which we usually want a logic. Because they do not correspond to closure operators, substructural logics cannot be used to determine the consequences of a collection of non-logical axioms like ZFC.

This is not to say that *every* logic must have a consequence relation that corresponds to a closure operator. There is nothing conceptually incoherent about a consequence relation that does not correspond to a closure operator. Non-Tarskian consequence relations are perfectly coherent. But any such consequence relation cannot be used to expand a theory in this way. Given that we at least sometimes want to use a logic in order to determine the consequences of a collection of axioms, we at least sometimes need to have a logic that corresponds to a closure operator.

Classical logic, paracomplete logics, and paraconsistent logics all do this: they each correspond to a closure operator, and allow us to use the logic to specify a single theory consisting of *the consequences* of those axioms. But substructural logics do not. For situations in which we need a closure operator, substructural logics cannot help us.

This suggests that substructural approaches are not really solving the paradoxes in all of the contexts in which we care about them, as much as they are simply changing the subject.

Quine [9] and Slater [14] objected to non-classical approaches to semantic paradox on the grounds that these approaches did not solve the paradoxes, but merely changed the subject. According to this objection, the liar paradox is formulated with a negation that obeys the classical negation rules. Introducing a logic with a negation that obeys different, weaker rules is simply changing languages, not solving the paradox as formulated in the original language.

Non-classical logicians have discussed and answered this objection.[2] But there is a similar, and I think deeper, version of the objection to be raised against the substructural approaches.

Both the nontransitive and the noncontractive approaches to paradox propose that there is something wrong with *cumulative reasoning*: one cannot safely start with a collection of premises, reason towards conclusions, add those conclusions to their collection, and repeat the process. Cumulative reasoning requires Cautious Cut, and Cautious Cut is ruled out by both nontransitive and noncontractive approaches.[3] But the fact of the matter is that much of what we want to do—in mathematics, in philosophy, and in our daily lives—requires cumulative reasoning.[4]

---

[2]See especially chapters 4 and 5 of [6].

[3]See, e.g., Ripley: "So neither the noncontractivist nor the nontransitivist can fully approve of cumulative reasoning. They diagnose different problems with it, and quarrel with different instances, but to the extent that cumulative reasoning is nonnegotiable, both noncontractive and nontransitive approaches are simply ruled out." [10].

[4]It does not help the situation to say that cumulative reasoning only fails in certain specific circumstances. Kripke [4] demonstrated quite convincingly that paradoxes can arise in our everyday lives, and so taking the substructural approach seriously would require giving up on cumulative reasoning *in general*.

The nontransitive and noncontractive approaches provide powerful tools that may serve us well in cases in which cumulative reasoning is not needed. These logics can provide perfectly useful consequence relations, that may apply well to many purposes. But the question nevertheless remains: when we need to reason cumulatively, what inferences are safe for us to make? To that question, the substructural approaches provide no answer. They simply change the subject.

The three papers presented here aim to illustrate some of the ways in which the noncontractive and nontransitive approaches are simply changing the subject. These logics can only solve the paradoxes as they apply to very particular, specialized notions of validity. But they do not solve paradoxes in full generality, and they do not solve the paradoxes as they apply to the notions of validity that are most important to us: the notions that allow for cumulative reasoning, that allow for the preservation of truth and justification and other desirable properties, and that tell us the consequences of the premises and axioms that we accept.

In the first paper, I argue that one of the essential features of a logic is the consequences that the logic provides for a collection of non-logical axioms. I show that two logics that have all the same validities and antivalidities (to be defined in the paper) at every inferential level can still give different consequences for the same axioms. I argue that one of the key reasons that we want a logic is to discover the consequences of a collection of axioms or assumptions. I argue that this poses a problem for nontransitive logics that go nontransitive "all the way up".

In the second paper, I present infinitary paradoxes that cause problems even in logics without the structural rule of contraction. These paradoxes involve infinitely many premises, and thus Contraction is in some sense "built in" and cannot be avoided. I argue that this is a symptom of a general phenomenon: noncontractive logics cannot be used to discover the consequences of a collection of axioms or assumptions. I argue that this is a problem for noncontractive approaches, as it means that we cannot use them to determine the consequences of collections of non-logical axioms like ZFC.

In the third paper, I introduce a metainferential hierarchy of validity curry paradoxes. I show that the validity curry paradox reappears at every level of inference, using metainferential versions of the rules involved in the original validity curry paradox. I argue that, if we take a uniform solution to the validity curry paradoxes to mean blocking the same step of the argument at every level, which is the notion of uniformity to defend substructural approaches, then no uniform solution is appealing. In particular, the nontransitive and noncontractive approaches must go nontransitive or noncontractive "all the way up", which is worse than giving up either rule at the first level alone. I then propose a solution on which we keep all of the usual structural rules, but block different validity rules at different levels. I argue that this is nevertheless a uniform solution.

The papers presented here are meant to be self-contained, and can be read in any order. The first paper has been published in the *Journal of Philosophical Logic*; the third paper is an expanded version of a paper presented to the Buenos Aires Logic Group's Work In Progress Seminar, and to the CUNY Graduate Center's Logic and Metaphysics Workshop. The papers can be read in any order, though each paper makes brief references to the papers that come before them.

Bibliography

[1] Barrio, Eduardo; Pailos, Federico; Szmuc, Damian (2021) "(Meta)inferential Levels of Entailment Beyond the Tarskian Paradigm" *Synthese* 198: 5265-5289.

[2] Beall, JC (2015) "Free of Detachment: Logic, Rationality, and Gluts" *Nous* 49(2): 410-23.

[3] Cintula, Petr and Paoli, Francesco (2016). Is Multiset Consequence Trivial? *Synthese*. https://doi-org.ezproxy.gc.cuny.edu/10.1007/s11229-016-1209-7.

[4] Kripke, Saul (1975) "Outline of a Theory of Truth" *Journal of Philosophy*, 72: 690-716.

[5] Mares, Edwin and Paoli, Francesco (2014) "Logical Consequence and the Paradoxes" *Journal of Philosophical Logic*, 43 (2-3), 439-469.

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

[7] Priest, Graham (1994) "The Structure of the Paradoxes of Self-Reference" *Mind*, 103: 25-41.

[8] Priest, Graham (1991) "Minimally Inconsistent LP" *Studia Logica*, 50(1): 321-31.

[9] Quine, Willard Van Orman (1970) *Philosophy of Logic*. Prentice-Hall.

[10] Ripley, David (2015a) "Comparing Substructural Theories of Truth" *Ergo*, 2, 299-328.

[11] Ripley, David (2015b) "Contraction and Closure" *Thought*, 4 (2), 131-138.

[12] Ripley, David (2013) "Paradoxes and Failures of Cut" *Australasian Journal of Philosophy* 91 (1): 139-164.

[13] Ripley, D. (2012) "Conservatively Extending Classical Logic with Transparent Truth" *Review of Symbolic Logic*, 5 (2): 354-378.

[14] Slater, Hartley (1995) "Paraconsistent Logics?" *Journal of Philosophical Logic*, 24, 451-454.

[15] Tarski, Alfred (1983) "On the Concept of Logical Consequence" In J. Corcoran, editor, *Logic, Semantics, Metamathematics: Papers from 1923 to 1938*, page 409-420. Hackett, Indianapolis, IN., 2nd edition.

Chapter 2

Supervaluations and the Strict-Tolerant Hierarchy

2.1 Introduction

There are several different logics available: classical logic, intuitionistic logic, the Strong Kleene logic K3, Priest's Logic of Paradox LP, supervaluationism, subvaluationism, etc. These logics are uncontroversially distinct: no one is under the impression that LP and intuitionistic logic are really two ways of presenting one and the same logic. However, the same logic can be formulated in different ways. Classical logic, for example, can be formulated in a natural deduction system, or a Hilbert-style calculus, or a multiple-conclusion sequent calculus. Classical logic can also be given a variety of semantics, and can be given several different axiomatizations. Yet all of these, intuitively, are simply different ways of presenting the same logic.

How can we identify a logic? Given two systems, how can we determine whether they are distinct logics, or are merely different presentations of the same logic? This is harder than it first seems. As a first pass, we might be tempted to say that a logic should be identified by its axioms and rules of inference. But this is too strict: it would count different formulations of classical logic as distinct logics, when in fact they are just different presentations of the same logic.

We might say that a logic should be identified by its set of theorems: the sentences that are derivable in the proof theory or get designated values at all models in the semantics. But this is too lax; classical logic, supervaluationism and Priest's paraconsistent logic LP all have the same theorems, and yet are uncontroversially distinct logics. One reason they are distinct is that they validate different inferences. For example, $A, A \to B \Rightarrow B$ is valid in classical logic and supervaluationism, but not in LP, while $A \vee B \Rightarrow A, B$ is valid in classical logic and LP but not in supervaluationism.[1]

We might therefore try to identify a logic by looking at its set of valid inferences. However,

[1] I specifically have in mind here supervaluationism with what Williamson [41] calls the *global* consequence relation.

this is also too lax. Two logics can agree on which inferences are valid, but disagree on which *meta*-inferences are valid. The logic ST, for example, introduced by Cobreros, Egré, Ripley and van Rooj (hereafter CERvR) in [8] and [9], has exactly the valid inferences of classical logic, but does not validate the meta-inferential Cut rule. It is therefore distinct from classical logic, which validates every instance of Cut.

We therefore must at least look at which metainferences are valid. But even the inferences and metainferences of a logic are not enough to identify a logic. In a recent paper, Barrio, Pailos and Szmuc (hereafter BPS) show that there are logics that have exactly the validities of classical logic up to arbitrarily high inference levels (in a sense to be made precise below), but then differ from classical logic after that [2]. Fitting [18] shows that this generalizes beyond classical logic; for K3, LP, FDE, and more, there are logics that have the exact same validities up to arbitrarily high levels of inference, but then differ at higher levels. BPS argue that this means that a logic must by identified by its valid inferences at *every* inferential level.

However, even this will not do the job. In a recent paper, Scambler builds on BPS's result, and shows that there are logics that have exactly the validities of classical logic at every level, and yet do not have the antivalidities of classical logic at any level (where an antivalidity is an inference such that every model is a counterexample to that inference) [36]. Fitting [17] shows that this also generalizes; there are logics that have the same validities as K3 or LP or FDE, yet do not have the same antivalidities. Scambler argues that this means that for two systems L and L' to be the same logic, they must at least have all of the same validities and antivalidities at every inferential level.

In this paper, I argue that this too is insufficient; there is a logic that has exactly the same validities and exactly the same antivalidities as classical logic at every inferential level, and yet is still intuitively distinct from classical logic.[2] In section 2.2 I introduce the logical

[2] I leave the question of whether this result generalizes beyond classical logic for future work.

framework that I'll be using in the rest of the paper. In section 2.3 I apply BPS's methods in a super/sub-valuationist framework, to construct notions of validity for inferences of every inferential level that have exactly the validities and antivalidities of classical logic. In section 2.4 I argue that the resulting logic is distinct from classical logic. In section 2.5 I discuss what this means for the problem of identifying a logic. In section 2.6 I consider whether this logic is paraconsistent, and what this means for ST and other logics with similar consequence relations. In section 2.7 I close with some concluding remarks.

2.2 Background

In this section I will introduce the logical framework that I'll be using in the rest of the paper. The formal system I present here requires a higher-order sequent calculus, in which we can have sequents for which both the premises and conclusions are themselves sequents of lower levels. Most of the notation used here comes from Scambler [36]. Most of the technical machinery for the slice hierarchy comes from [2] and [28], although higher-order sequent systems were earlier explored by Kosta Došen, for example in [11], [12], [13], and [14]. However, I will be extending the hierarchy into the transfinite.[3] The machinery for a super/sub-valuationist mixed consequence relation comes from [10].

2.2.1 Languages

We define our set of languages inductively over the ordinals:

Base Case: Let $\mathcal{L}_0$ be a standard propositional language, with propositional constants p_i and the connectives $\neg, \wedge, \vee$.

[3] Scambler [37] also extends the slice hierarchies into the transfinite, though his methods for doing so are slightly different than mine. I suspect that the results using my transfinite framework carry over to his, and vice-versa, but have not confirmed this.

Successor Case: Given a language L_α for some ordinal α, let $\mathcal{L}_{\alpha+1}$ be the set of all pairs of sets $\langle \Gamma, \Delta \rangle$ such that $\Gamma \cup \Delta \subseteq \mathcal{L}_\alpha$. I'll write $\langle \Gamma, \Delta \rangle$ in a sequent format, as $\Gamma \Rightarrow_{\alpha+1} \Delta$. We'll call $\Gamma \Rightarrow_\alpha \Delta$ an "inference of order α" or an "α-inference". So $p \wedge q \Rightarrow_1 p$ is a 1-inference, $\{p \Rightarrow_1 p \wedge q\} \Rightarrow_2 \{\Rightarrow_1 q\}$ is a 2-inference, and so on. I will omit set brackets when no confusion can result.

Limit Case: Given languages L_β for all $\beta < \lambda$, let $\mathcal{L}_\lambda$ be the set of all pairs of sets $\langle \Gamma, \Delta \rangle$ such that $\Gamma \cup \Delta \subseteq \bigcup_{\beta<\lambda} L_\beta$. We'll call $\Gamma \Rightarrow_\lambda \Delta$ an "inference of order λ" or a "λ-inference". I will again will omit set brackets when no confusion can result.

Note that the limit-ordinal languages will be cumulative, in the terminology of [37]: the sequents of limit-ordinal languages can contain inferences from *any* lower level. The successor-ordinal languages, however, will not be cumulative in this sense: $\alpha + 1$-sequents can only contain α-inferences.

2.2.2 Valuations and Models

Let a *Boolean valuation* be a function $v : \mathcal{L}_0 \to \{0, 1\}$ such that the connectives $\neg, \wedge, \vee$ obey the truth tables of classical logic. The set of Boolean valuations is the set of models for propositional classical logic. I'll call propositional classical logic CL.

Let a *supervaluation* (SV) model be a nonempty set of Boolean valuations. I'll use $\mathfrak{V}$ to denote the set of SV models. These models are based on the formal semantics for supervaluationist and subvaluationist logics.[4]

[4]For a general overview of supervaluationism, see [23]; for earlier presentations and defenses see [15], [24], [27], and [38]; see [39] for a discussion of different consequence relations compatible with supervaluationist semantics. For a general overview of subvaluationism, see [6]; for defenses and earlier presentations, see [7], [20], [21], [22], and [40].

2.2.3 Notions of Validity

We say a *notion of α-validity* is a function $V : \mathfrak{V} \times \mathcal{L}_\alpha \to \{0,1\}$. We say a model $m \in \mathfrak{V}$ *satisfies* an α-inference $\Gamma \Rightarrow_\alpha \Delta$ iff $V(m, \Gamma \Rightarrow_\alpha \Delta) = 1$. We can think of a notion of 0-validity as a notion of "truth-in-a-model" that determines which sentences get designated values at each model. A notion of $\alpha + 1$-validity tells us which inferences between sets of α-inferences are valid, and a notion of λ-inference for limit ordinal λ tells us which inferences between sets of lower-level inferences are valid.

I will write $V \vDash_m \Gamma \Rightarrow_\alpha \Delta$ in place of $V(m, \Gamma \Rightarrow_\alpha \Delta) = 1$, and $V \nvDash_m \Gamma \Rightarrow_\alpha \Delta$ in place of $V(m, \Gamma \Rightarrow_\alpha \Delta) = 0$.

When no confusion can result, I will sometimes use Φ in place of $\Gamma \Rightarrow_\alpha \Delta$ for convenience. When it is necessary to indicate that Φ is an α-inference for some ordinal α, I will write Φ as Φ_α.

We say that an inference Φ is *valid* on a notion of α-validity iff for all $m \in \mathfrak{V}$, $V \vDash_m \Phi$. We will write this as $V \vDash \Phi$.

We say that an inference Φ is *anti-valid* on a notion of α-validity iff for all $m \in \mathfrak{V}$, $V \nvDash_m \Phi$. We will write this as $V \dashv \Phi$.

In other words, an inference is valid iff no model is a counterexample; an inference is antivalid if every model is a counterexample.

For all α and all $\Phi \in \mathcal{L}_\alpha$, we say that $CL \vDash \Phi$ whenever Φ is a valid α-inference in classical logic CL.

2.2.4 Slice Hierarchies

BPS and Scambler use the following definition to produce logics of arbitrarily high (finite) inference levels with notions of α-validity:

Definition 2.2.1. ***Successor Slice*** *Let V and U be notions of α-validity. Then the slice* *of V and U, which we write as V/U, is the notion of $\alpha+1$ validity such that for all $m \in \mathfrak{V}$: $V/U \vDash_m \Gamma \Rightarrow_{\alpha+1} \Delta$ iff $(\exists\gamma \in \Gamma)V \nvDash_m \gamma$ or $(\exists\delta \in \Delta)U \vDash_m \delta$.*

We can extend this to the transfinite by adding the limit case:

Definition 2.2.2. ***Limit Slice**: Let $\{V_\beta\}_{\beta<\lambda}$ and $\{U_\beta\}_{\beta<\lambda}$ be sets of notions of validity for all $\beta < \lambda$. Then the limit slice of $\{V_\beta\}_{\beta<\lambda}$ and $\{U_\beta\}_{\beta<\lambda}$, which we write as $V_{<\lambda}/U_{<\lambda}$ is the notion of λ-validity such that for all $m \in \mathfrak{V}$: $V_{<\lambda}/U_{<\lambda} \vDash_m \Gamma \Rightarrow_\lambda \Delta$ iff either:*

$\exists\gamma \in \Gamma$ such that γ is a β-inference for some $\beta < \lambda$ and $V_\beta \nvDash_m \gamma$, or

$\exists\delta \in \Delta$ such that δ is a β-inference for some $\beta < \lambda$ and $U_\beta \vDash_m \delta$.

Given two notions of α-validity V and U, we can build the transfinite slice hierarchy over V and U:

$$V, U$$

$$V/V,\ V/U,\ U/V,\ U/U$$

$$\vdots$$

Using slices like the ones defined above, BPS show that one can use Strong Kleene 3-valued models to build a slice hierarchy based on CERvR's Strict-Tolerant logic ST. They prove that for every inferential level $n < \omega$, there is a logic that has exactly the validities of classical logic up to order n, but differs from classical logic at higher inferences levels. They suggest that classical logic therefore must be identified by its valid inferences at *every* (finite) inferential level. In [18], Fitting shows that this can be generalized beyond classical logic. It therefore seems that *no* logic can be characterized by its valid inferences; we must look at its valid metainferences at every metainferential level.

However, in [36] and [37], Scambler shows that even a logic with exactly the validities of classical logic at every inferential level can differ from classical logic. He demonstrates that what he calls the "tolerant twist logic" has exactly the validities of classical logic at every inferential level, but unlike classical logic, has no antivalidities. In [17] and [16], Fitting shows that this too generalizes beyond classical logic. For any logic, we cannot characterize that logic just by looking at the validities of every level. We must at least look at the validities *and antivalidities* of every inferential level.

In what follows, I construct a transfinite slice hierarchy that gives us a logic with exactly the validities of classical logic at every inferential level, and exactly the antivalidities of classical logic at every inferential level. I will then argue that this logic still should not be identified with classical logic, and that we therefore need a new criterion for identifying logics.

2.2.5 A Note about Local and Global Validity

It is important to note that the definitions of successor and limit slices given above will produce notions of *local* validity, rather than notions of *global* validity. When dealing with metainferences, there are at least two ways to define metainferential validity over a class of models: *local* validity and *global* validity.[5]

In traditional (unsliced) contexts, local validity can be thought of as preservation of satisfaction, while global validity can be thought of as preservation of validity. A metainference is locally valid iff at every model, either some conclusion inference is satisfied or some premise inference is not. A metainference is globally valid iff either some conclusion inference in valid or some premise inference is valid. For example, the metainference $\{p \Rightarrow_1 q\} \Rightarrow_2 \{p \Rightarrow_1 q \wedge r\}$ is globally valid in classical logic, because the premise inference $p \Rightarrow_1 q$ is not valid. But it is not locally valid in classical logic, because there are classical valuations at which $p \Rightarrow_1 q$ is satisfied but $p \Rightarrow_1 q \wedge r$ is not. In general, local validity implies global validity, but global validity does not imply local validity.

[5]For a detailed discussion of the distinction between local and global metainferential validity, see [19].

We can generalize these notions to slice notions of validity: an inference is V/U locally valid iff at every model, either some premise is not V-satisfied or some conclusion is U-satisfied. Similarly, an inference is V/U globally valid iff either some premise is not V-valid or some conclusion is U-valid.

The notions of validity generated by the above definitions for successor and limit slices are local, rather than global, notions of validity. Following [2], [28], and [36], I will be restricting attention to *local* validity at each level, rather than global validity. I will therefore use $CL \vDash \Phi$ whenever Φ is a locally valid inference in classical logic. There is a sense in which local validity is more fine-grained than global validity: any two logics with the same locally valid inferences must have the same globally valid inferences, but not vice-versa. ST and classical logic have the same globally valid 2-inferences (in the empty signature), but have different locally valid inferences: Cut is locally valid in classical logic, but is not locally valid in ST (even in the empty signature). Therefore, when we see in the following sections that the resulting logic has exactly the local validities of classical logic, it immediately follows that it also has exactly the global validities of classical logic.

2.3 The LM Hierarchy

2.3.1 Two Notions of 0-validity

There are at least two interesting notions of 0-validity that can be defined over the set of SV models $\mathfrak{V}$. These correspond to the notions of 0-validity for supervaluationism and subvaluationism, so I'll call them P (for suPervaluationism) and B (for suBvaluationism):

$P \vDash_m \phi$ iff $v(\phi) = 1$ for all $v \in m$

$B \vDash_m \phi$ iff $v(\phi) = 1$ for some $v \in m$

It is easy to see that for all 0-inferences, $P \vDash \phi$ iff $B \vDash \phi$ iff $CL \vDash \phi$. If every Boolean valuation assigns 1 to ϕ, then every set of Boolean valuations contains only valuations that assign 1 to ϕ. It is also easy to see that $P \Dashv \phi$ iff $B \Dashv \phi$ iff $CL \Dashv \phi$. If every Boolean valuation assigns 0 to ϕ, then no set of Boolean valuations has any valuations that assign 1 to ϕ.

2.3.2 Six Notions of 1-validity

We could slice together P and B to form a hierarchy of notions of validity. For example, there are four notions of 1-validity that we can get just by slicing P and B:

$P/P \vDash_m \Gamma \Rightarrow_1 \Delta$ iff either $(\exists \gamma \in \Gamma)(\exists v \in m)v(\gamma) = 0$ or $(\exists \delta \in \Delta)(\forall v \in m)v(\delta) = 1$

$B/B \vDash_m \Gamma \Rightarrow_1 \Delta$ iff either $(\exists \gamma \in \Gamma)(\forall v \in m)v(\gamma) = 0$ or $(\exists \delta \in \Delta)(\exists v \in m)v(\delta) = 1$

$B/P \vDash_m \Gamma \Rightarrow_1 \Delta$ iff either $(\exists \gamma \in \Gamma)(\forall v \in m)v(\gamma) = 0$ or $(\exists \delta \in \Delta)(\forall v \in m)v(\delta) = 1$

$P/B \vDash_m \Gamma \Rightarrow_1 \Delta$ iff either $(\exists \gamma \in \Gamma)(\exists v \in m)v(\gamma) = 0$ or $(\exists \delta \in \Delta)(\exists v \in m)v(\delta) = 1$

P/P and B/B correspond to the (global) 1-validity consequence relations for supervaluationism and subvaluationism, respectively. Neither has all of the 1-validities of classical logic. For example, $P/P \nvDash \{\} \Rightarrow_1 A, \neg A$, and $B/B \nvDash A, \neg A \Rightarrow_1 \{\}$.

The mixed-condition 1-validity consequence relations B/P and P/B are effectively the super/sub-valuation equivalents of the K3/LP consequence relations TS and ST.[6] Like TS, B/P is not reflexive, in the sense that $A \Rightarrow_1 A$ is not valid unless A is a classical tautology or classical contradiction. Like ST, P/B has exactly the same valid 1-inferences as classical logic. And like ST, there is a sense in which P/B is not transitive. In particular, $P/B \vDash_m \Gamma \Rightarrow_1 A$ and $P/B \vDash_m A \Rightarrow_1 \Delta$ do not imply $P/B \vDash_m \Gamma \Rightarrow_1 \Delta$.[7] Although ST and P/B both have exactly the 1-validities of classical logic, P/B has an additional

[6] For more information about the strict-tolerant approach in a super/sub-valuationist setting, see [10].

[7] There are many different ways to use the term "transitive" when discussing consequence relations. See [31] for a survey of several different notions. My use of the term is somewhat idiosyncratic, in that it is really transitivity of *satisfaction* that I have in mind, rather than transitivity of *validity*. However, I take this to be an important and distinctive feature of the logic: it means that the set of sentences that are true in a model are not closed under the valid inferences of the logic. This has significant consequences for how the logic handles non-logical axioms, which will be discussed more in sections 2.4 and 2.6.

feature that ST does not have: P/B has all of the 1-validities of classical logic *and* all of the 1-antivalidities of classical logic.

B/P and P/B could be sliced together to construct a notion of 2-validity, which will again have all of the 2-validities of classical logic. However, this notion of 2-validity will only have *some* of the antivalidities of classical logic. For example, the 2-inference from $A \Rightarrow_1 A$ to $\Rightarrow_1 A \wedge \neg A$ is classically antivalid, but it is not antivalid in BP/PB: there are SV models at which $A \Rightarrow_1 A$ is not B/P-satisfied. In order to construct a logic with all of validities and all of the antivalidities of classical logic, we need to build our hierarchy using different notions of 1-validity.

Instead of slicing together these two notions of 1-validity, we can define notions of 1-validity directly, and build the hierarchy out of those. Our first such notion of 1-validity, L, corresponds to what Williamson [41] calls the "local" consequence relation on supervaluation models:

$$L \vDash_m \Gamma \Rightarrow_1 \Delta \text{ iff } (\forall v \in m)[(\exists \gamma \in \Gamma)v(\gamma) = 0 \vee (\exists \delta \in \Delta)v(\delta) = 1]$$

A 1-inference is L-satisfied at a model iff it is classically satisfied at every valuation in the model. It's easy to see that this notion of 1-validity has exactly the 1-validities and 1-antivalidities of classical logic.

In addition to L, we can also define a more tolerant consequence relation M:

$$M \vDash_m \Gamma \Rightarrow_1 \Delta \text{ iff } (\exists v \in m)[(\exists \gamma \in \Gamma)v(\gamma) = 0 \vee (\exists \delta \in \Delta)v(\delta) = 1]$$

A 1-inference is M-satisfied at a model iff it is classically satisfied at *some* valuation in the model. M is, in fact, equivalent to P/B.[8] Thus, like L, M has exactly the 1-validities and 1-antivalidities of classical logic. But despite this, M (or equivalently, P/B) is not transitive.

[8]This is due to the fact that the existential quantifier distributes over disjunction in our classical meta-language.

To see this, let m be a model containing two valuations u and v, such that $v(p) = 1$, $v(q) = 0$, $u(p) = 0$, and $u(q) = 0$. In this case $M \vDash_m \Rightarrow_1 p$ and $M \vDash_m p \Rightarrow_1 q$, and yet $M \nvDash_m \Rightarrow_1 q$.

There is one fact about L and M that will be important in constructing the LM hierarchy: for any SV model $m \in \mathfrak{V}$, $L \vDash_m \{\} \Rightarrow_1 \phi$ iff $P \vDash_m \phi$, and $M \vDash \{\} \Rightarrow_1 \phi$ iff $B \vDash \phi$. Without this property, we couldn't use these notions of 0- and 1-validity together to construct anything that deserved to be called a "logic".

It is easy to see, but important to note before we continue, that all six of these notions of 1-validity have the following feature: given a singleton model $\{v\}$, each notion of validity gives us $\vDash_{\{v\}} \Gamma \Rightarrow_1 \Delta$ iff $CL \vDash_v \Gamma \Rightarrow_1 \Delta$. Notions of α-validity at any level can have the equivalent property for α-inferences. I will call this the *singleton property*:

Definition 2.3.1. ***Singleton Property***

A notion of α-validity V has the singleton property if for all singleton models $\{v\}$ and all α-inferences $\Gamma \Rightarrow_\alpha \Delta$, $V \vDash_{\{v\}} \Gamma \Rightarrow_\alpha \Delta$ iff $CL \vDash_v \Gamma \Rightarrow_\alpha \Delta$.

Note that P and B, our notions of 0-validity, also have the singleton property. We can now use P, B, L and M to construct a hierarchy of notions of validity that have all of the validities and antivalidities of classical logic at every level of inference.

2.3.3 The Hierarchy

We will define what I will call the LM hierarchy inductively:

Base cases:

L_0 = P, M_0 = B

L_1 = L, M_1 = M

Successor Case: for $\alpha \geqslant 1$:

$L_{\alpha+1} = M_\alpha / L_\alpha$, $M_{\alpha+1} = L_\alpha / M_\alpha$

Limit Case: for limit ordinals λ:

$L_\lambda = M_{<\lambda}/L_{<\lambda}$, $M_\lambda = L_{<\lambda}/M_{<\lambda}$

Call M_∞ the logic that evaluates α-validity at every ordinal α in accordance with M_α. Similarly for L_∞.

My primary goal in this section will be to show that M_∞ has exactly the validities and antivalidities of classical logic at every inference level. First, we need to show that every notion of validity in the hierarchy has the singleton property:

Lemma 1. *If two notions of α-validity X_α and Y_α both have the singleton property, then their slice X_α/Y_α also has the singleton property.*

Proof. $CL \vDash_v \Gamma \Rightarrow_{\alpha+1} \Delta$ iff either there is a $\gamma \in \Gamma$ such that $CL \nvDash_v \gamma$, or there is a $\delta \in \Delta$ such that $CL \vDash_v \delta$. Since X_α and Y_α both have the singleton property, $CL \nvDash_v \gamma$ iff $X_\alpha \nvDash_{\{v\}} \gamma$ and $CL \vDash_v \delta$ iff $Y_\alpha \vDash_{\{v\}} \delta$. $X_\alpha/Y_\alpha \vDash_{\{v\}} \Gamma \Rightarrow_{\alpha+1} \Delta$ iff there is a $\gamma \in \Gamma$ such that $X_\alpha \nvDash_{\{v\}} \gamma$ or there is a $\delta \in \Delta$ such that $Y_\alpha \vDash_{\{v\}} \delta$. Therefore $CL \vDash_v \Gamma \Rightarrow_{\alpha+1} \Delta$ iff $X_\alpha/Y_\alpha \vDash_{\{v\}} \Gamma \Rightarrow_{\alpha+1} \Delta$. □

This also holds for the limit slices:

Lemma 2. *If every notion of validity $X_i \in \{X_\beta\}_{\beta<\lambda}$ and $Y_i \in \{Y_\beta\}_{\beta<\lambda}$ has the singleton property, then $X_{<\lambda}/Y_{<\lambda}$ has the singleton property.*

Proof. Suppose $CL \vDash_v \Gamma \Rightarrow_\lambda \Delta$. Then either $CL \nvDash_v \gamma$ for some $\gamma \in \Gamma$, or $CL \vDash_v \delta$ for some $\delta \in \Delta$. By IH, for all $\beta < \lambda$, $X_\beta \vDash_{\{v\}} \Phi_\beta$ iff $CL \vDash_v \Phi_\beta$, and $Y_\beta \vDash_{\{v\}} \Phi_\beta$ iff $CL \vDash_v \Phi_\beta$. If $CL \nvDash_v \gamma_\beta$ for some $\gamma \in \Gamma$ and $\beta < \lambda$, then $X_\beta \nvDash_{\{v\}} \gamma_\beta$, and so $X_{<\lambda}/Y_{<\lambda} \vDash_{\{v\}} \Gamma \Rightarrow_\lambda \Delta$. Otherwise, $CL \vDash_v \delta_\beta$ for some $\delta \in \Delta$ and some $\beta < \lambda$. Then $Y_\beta \vDash_{\{v\}} \delta_\beta$, and so $X_{<\lambda}/Y_{<\lambda} \vDash_{\{v\}} \Gamma \Rightarrow_\lambda \Delta$.

The reverse direction follows the same pattern. □

It follows that if two notions of validity V and U have the singleton property, then every notion of validity in the transfinite hierarchy over V and U has the singleton property. Since

P, B, L, and M all have the singleton property, this means that every notion of validity in M_∞ and in L_∞ has the singleton property. We can use this fact to prove sufficient conditions for a slice notion of $\alpha + 1$-validity to have all of the validities and all of the antivalidities of classical logic.

Lemma 3. *Let X_α and Y_α be two notions of α-validity defined over the set of SV models $\mathfrak{V}$ that both have the singleton property. Suppose that for all SV models m, $X_\alpha \vDash_m \Phi$ implies $CL \vDash_v \Phi$ for all $v \in m$, and $Y_\alpha \nvDash_m \Phi$ implies $CL \nvDash_v \Phi$ for all $v \in m$. Then $X_\alpha/Y_\alpha \vDash \Gamma \Rightarrow_{\alpha+1} \Delta$ iff $CL \vDash \Gamma \Rightarrow_{\alpha+1} \Delta$.*

Proof. The left-to-right direction follows from lemma 1. Since X_α and Y_α both have the singleton property, X_α/Y_α does too. So $CL \nvDash \Phi$ implies that $X_\alpha/Y_\alpha \nvDash \Phi$: if some valuation v is a CL counterexample to Φ, then $\{v\}$ will be a X_α/Y_α counterexample to Φ.

For the right-to-left direction, suppose $X_\alpha/Y_\alpha \nvDash \Gamma \Rightarrow_{\alpha+1} \Delta$. Then there is an SV model $m \in \mathfrak{V}$ such that $X_\alpha/Y_\alpha \nvDash_m \Gamma \Rightarrow_{\alpha+1} \Delta$. It follows that $X_\alpha \vDash_m \gamma$ for all $\gamma \in \Gamma$ and $Y_\alpha \nvDash_m \delta$ for all $\delta \in \Delta$. By assumption, $X_\alpha \vDash_m \Phi$ implies $CL \vDash_v \Phi$ for all $v \in m$, and $Y_\alpha \nvDash_m \Phi$ implies $CL \nvDash_v \Phi$ for all $v \in m$. Therefore every $v \in m$ is such that $CL \vDash_v \gamma$ for all γ and $CL \nvDash_v \delta$ for all $\delta \in \Delta$. Therefore $CL \nvDash \Gamma \Rightarrow_{\alpha+1} \Delta$. □

Lemma 4. *Let X_α and Y_α be two notions of α-validity defined over the set of SV models $\mathfrak{V}$ that both have the singleton property. Suppose that for all SV models $m \in \mathfrak{V}$, $CL \vDash_v \Phi$ for all $v \in m$ implies $X_\alpha \vDash_m \Phi$, and $CL \nvDash_v \Phi$ for all $v \in m$ implies $Y_\alpha \nvDash_m \Phi$. Then $X_\alpha/Y_\alpha \Dashv \Gamma \Rightarrow_{\alpha+1} \Delta$ iff $CL \Dashv \Gamma \Rightarrow_{\alpha+1} \Delta$.*

Proof. The left-to-right direction follows from lemma 1. Since X_α and Y_α both have the singleton property, X_α/Y_α does too. So $CL \not\Dashv \Phi$ implies that $X_\alpha/Y_\alpha \not\Dashv \Phi$: if some valuation v is not a CL counterexample to Φ, then $\{v\}$ will not be a X_α/Y_α counterexample to Φ.

For the right-to-left direction, suppose $CL \Dashv \Gamma \Rightarrow_{\alpha+1} \Delta$. Then for every Boolean valuation v, $CL \vDash_v \gamma$ for all $\gamma \in \Gamma$ and $CL \nvDash \delta$ for all $\delta \in \Delta$. By assumption, $CL \vDash_v \Phi$ for all

$v \in m$ implies $X_\alpha \vDash_m \Phi$, and $CL \nvDash_v \Phi$ for all $v \in m$ implies $Y_\alpha \nvDash_m \Phi$. It follows that for every SV model $m \in \mathfrak{V}$, $X_\alpha \vDash_m \Phi$ and $Y_\alpha \nvDash_m \Phi$. Therefore, $X_\alpha/Y_\alpha \dashv \Gamma \Rightarrow_{\alpha+1} \Delta$. □

These conditions also hold for the limit slices:

Lemma 5. *Let $\{X_\beta\}_{\beta<\lambda}$ and $\{Y_\beta\}_{\beta<\lambda}$ be two sets of notions of β-validity defined over the set of SV models $\mathfrak{V}$ for all $\beta < \lambda$ that all have the singleton property. Suppose that for all $\beta < \lambda$ and all SV models $m \in \mathfrak{V}$, $X_\beta \vDash_m \Phi$ implies $CL \vDash_v \Phi$ for all $v \in m$, and $Y_\beta \nvDash_m \Phi$ implies $CL \nvDash_v \Phi$ for all $v \in m$. Then $X_{<\lambda}/Y_{<\lambda} \vDash \Gamma \Rightarrow_\lambda \Delta$ iff $CL \vDash \Gamma \Rightarrow_\lambda \Delta$.*

Proof. The left-to-right direction follows from the singleton property.

For the right-to-left direction, suppose that $X_{<\lambda}/Y_{<\lambda} \nvDash \Gamma \Rightarrow_\lambda \Delta$. Then there is an SV model $m \in \mathfrak{V}$ such that for all $\gamma \in \Gamma$ and all $\delta \in \Delta$ and all $\beta < \lambda$, $X_\beta \vDash_m \gamma_\beta$ and $Y_\beta \nvDash_m \delta_\beta$. It follows that for all $v \in m$, $CL \vDash_v \gamma$ and $CL \nvDash_v \delta$ for all $\gamma \in \Gamma$ and $\delta \in \Delta$. Therefore, $CL \nvDash \Gamma \Rightarrow_\lambda \Delta$. □

Lemma 6. *Let $\{X_\beta\}_{\beta<\lambda}$ and $\{Y_\beta\}_{\beta<\lambda}$ be two sets of notions of β-validity defined over the set of SV models $\mathfrak{V}$ for all $\beta < \lambda$ that all have the singleton property. Suppose that for all $\beta < \lambda$, all β-inferences Φ, and all SV models $m \in \mathfrak{V}$, $CL \vDash_v \Phi$ for all $v \in m$ implies $X_\beta \vDash_m \Phi$, and $CL \nvDash_v \Phi$ for all $v \in m$ implies $Y_\beta \nvDash_m \Phi$. Then $X_{<\lambda}/Y_{<\lambda} \dashv \Gamma \Rightarrow_\lambda \Delta$ iff $CL \dashv \Gamma \Rightarrow_\lambda \Delta$.*

Proof. The left-to-right direction follows from the singleton property.

For the right-to-left direction, suppose that $CL \dashv \Gamma \Rightarrow_\lambda \Delta$. Then for every Boolean valuation v, $CL \vDash_v \gamma$ for all $\gamma \in \Gamma$, and $CL \nvDash_v \delta$ for all $\delta \in \Delta$. By assumption, $CL \vDash_v \Phi_\beta$ for all $v \in m$ implies $X_\beta \vDash_m \Phi_\beta$, and $CL \nvDash_v \Phi_\beta$ for all $v \in m$ implies $Y_\beta \nvDash_m \Phi_\beta$. It follows that for all SV models $m \in \mathfrak{V}$, $X_\beta \vDash_m \gamma_\beta$ and $Y_\beta \nvDash_m \delta_\beta$. Therefore $X_{<\lambda}/Y_{<\lambda} \dashv \Gamma \Rightarrow_\lambda \Delta$. □

So to show that the logic M_∞ has all of the validities and antivalidities of classical logic at every inferential level, it suffices to show that for all ordinals $\alpha \geqslant 1$ and all SV models $m \in \mathfrak{V}$, $L_\alpha \vDash_m \Phi$ iff $CL \vDash_v \Phi$ for all $v \in m$, and $M_\alpha \nvDash_m \Phi$ iff $CL \vDash_v \Phi$ for all $v \in m$.

Lemma 7. *For all* $\alpha \geqslant 1$, $L_\alpha \vDash_m \Phi$ *iff* $CL \vDash_v \Phi$ *for all* $v \in m$, *and* $M_\alpha \nvDash_m \Phi$ *iff* $CL \nvDash_v \Phi$ *for all* $v \in m$.

Proof. By induction. The base case is L and M, and follows immediately from the definition of L and M.

For the successor case, we take the inductive hypothesis that for all SV models $m \in \mathfrak{V}$ and all $\Phi \in \mathcal{L}_\alpha$, $L_\alpha \vDash_m \Phi$ iff $CL \vDash_v \Phi$ for all $v \in m$ and $M_\alpha \nvDash_m \Phi$ iff $CL \nvDash_v \Phi$ for all $v \in m$.

For $L_{\alpha+1}$, we note that $L_{\alpha+1} \vDash_m \Gamma \Rightarrow_{\alpha+1} \Delta$ iff either $(\exists\gamma \in \Gamma) M_\alpha \nvDash_m \gamma$ or $(\exists\delta \in \Delta) L_\alpha \vDash_m \delta$. By IH, for all $\gamma \in \Gamma$ and all $\delta \in \Delta$, $M_\alpha \nvDash_m \gamma$ iff $CL \nvDash_v \gamma$ for all $v \in m$, and $L_\alpha \vDash_m \delta$ iff $CL \vDash_v \delta$ for all $v \in m$. By definition, $(\exists\gamma \in \Gamma) CL \nvDash_v \gamma$ for all $v \in m$ or $(\exists\delta \in \Delta) CL \vDash_v \delta$ for all $v \in m$ iff $CL \vDash_v \Gamma \Rightarrow_{\alpha+1} \Delta$ for all $v \in m$.

For $M_{\alpha+1}$, we note that $M_{\alpha+1} \nvDash_m \Gamma \Rightarrow_{\alpha+1} \Delta$ iff $(\forall\gamma \in \Gamma) L_\alpha \vDash_m \gamma$ and $(\forall\delta \in \Delta) M_\alpha \nvDash_m \delta$. By IH, for all $\gamma \in \Gamma$ and all $\delta \in \Delta$, $L_\alpha \vDash_m \gamma$ iff $CL \vDash_v \gamma$ for all $v \in m$, and $M_\alpha \nvDash_m \delta$ iff $CL \nvDash_v \delta$ for all $v \in m$. By definition, $CL \vDash_v \gamma$ for all $v \in m$ and $CL \nvDash_v \delta$ for all $v \in m$ iff $CL \nvDash \Gamma \Rightarrow_{\alpha+1} \Delta$ for all $v \in m$.

For the limit case, we take the inductive hypothesis that for all $\beta < \lambda$, for all SV models $m \in \mathfrak{V}$ and for all $\Phi \in \bigcup_{\beta<\lambda} \mathcal{L}_\beta$, $L_\beta \vDash_m \Phi$ iff $CL \vDash_v \Phi$ for all $v \in m$ and $M_\beta \nvDash_m \Phi$ iff $CL \nvDash_v \Phi$ for all $v \in m$.

For L_λ, we note that $L_\lambda \vDash_m \Gamma \Rightarrow_\lambda \Delta$ iff either $(\exists\gamma \in \Gamma)(\exists\beta < \lambda) M_\beta \nvDash_m \gamma$ or $(\exists\delta \in \Delta)(\exists\beta < \lambda) L_\beta \vDash_m \delta$. By IH, for all $\gamma \in \Gamma$ and all $\delta \in \Delta$ and all $\beta < \lambda$, $M_\beta \nvDash_m \gamma_\beta$ iff $CL \nvDash_v \gamma_\beta$ for all $v \in m$, and $L_\beta \vDash_m \delta_\beta$ iff $CL \vDash_v \delta_\beta$ for all $v \in m$. By definition, $(\exists\gamma \in \Gamma) CL \nvDash_v \gamma$ for all $v \in m$ or $(\exists\delta \in \Delta) CL \vDash_v \delta$ for all $v \in m$ iff $CL \vDash_v \Gamma \Rightarrow_\lambda \Delta$ for all $v \in m$.

For M_λ, we note that $M_\lambda \nvDash_m \Gamma \Rightarrow_\lambda \Delta$ iff $(\forall\gamma \in \Gamma)(\exists\beta < \lambda) L_\beta \vDash_m \gamma$ and $(\forall\delta \in \Delta)(\exists\beta < \lambda) M_\beta \nvDash_m \delta$. By IH, for all $\gamma \in \Gamma$ and all $\delta \in \Delta$ and all $\beta < \lambda$, $L_\beta \vDash_m \gamma_\beta$ iff $CL \vDash_v \gamma_\beta$ for all $v \in m$, and $M_\beta \nvDash_m \delta_\beta$ iff $CL \nvDash_v \delta_\beta$ for all $v \in m$. By definition, $CL \vDash_v \gamma$ for all $v \in m$ and $CL \nvDash_v \delta$ for all $v \in m$ iff $CL \nvDash_v \Gamma \Rightarrow_\lambda \Delta$ for all $v \in m$. □

With this, we can now prove the primary result of this section:

Theorem 8. *For every ordinal α, $M_\infty \models \Phi_\alpha$ iff $CL \models \Phi_\alpha$, and $M_\infty \Dashv \Phi_\alpha$ iff $CL \Dashv \Phi_\alpha$.*

Proof. Follows immediately from lemmas 3, 4, 5, 6, 7, and the fact that B has exactly the 0-validities and 0-antivalidities of classical logic and M has exactly the 1-validities and 1-antivalidities of classical logic. □

M_∞ therefore has exactly the validities and antivalidities of classical logic at every inferential level.

2.4 Is M_∞ Classical Logic?

We have just shown that the logic M_∞ has exactly the validities and antivalidities of classical logic at every inferential level. One might therefore be tempted to identify M_∞ with classical logic. However, this would be a mistake.

To illustrate why, we need to look at how the two logics handle the addition of non-logical axioms. When presented with the same set of axioms in the same language, M_∞ and classical logic will generate different *theories*, at least in some cases. I take this to be a reason for thinking that the two logics are distinct.

We often want to use a logic to prove theorems from sets of axioms, as in the case of ZFC or Peano Arithmetic. This does not involve moving to a new logic; a single logic, like classical logic, can be used with a variety of theory-specific non-logical axioms and still be the same logic in each case. Any set of axioms will have certain consequences in a logic. Given a logic L and a set of axioms Γ, let the *theory* of Γ in L be the set of formulae that are true (satisfied) at all L-models at which the members of Γ are true (satisfied).[9]

[9]Although "theory" is defined here semantically, this definition is equivalent to the proof-theoretic notion of a theory as the set of formulae provable from a set of axioms. I use the semantic definition here for the simple reason that I do not currently have a proof theory for the model-theoretically-defined logic M_∞.

Suppose that we are presented with two logics, L and L′. Suppose that, given the same language $\mathcal{L}$ and the same set of non-logical axioms Γ, the theory generated by Γ in L is strictly greater than the theory generated by Γ in L′. In that case, I take it that L and L′ can safely be considered different logics. What follows from a given set of axioms does not and should not depend on how the logic is presented. It would be quite a shock to discover that whether or not the continuum hypothesis follows from ZFC depends on how we present first order classical logic: in one presentation, the continuum hypothesis is independent of the axioms of ZFC, but in another presentation it is a theorem. That simply cannot happen; there must be some determinate, presentation-independent fact of the matter as to what the consequences of such-and-such axioms are in a given logic. What follows from a set of axioms in a logic is not a matter of presentation; it is an essential feature of the logic.

With that in mind, let's consider how classical logic and M_∞ behave given the same set of axioms in the same language. In the standard propositional language used in the previous section, take p_1 and $\neg p_1$ as axioms in both classical logic and M_∞. In the sequent calculus presentation that we've been using, one way to do this is to add $\{\Rightarrow_1 p_1, \Rightarrow_1 \neg p_1\}$ as a set of axioms in our formal system.[10] We can then examine the theory of $\{\Rightarrow_1 p_1, \Rightarrow_1 \neg p_1\}$ in each logic.[11]

The theory generated by these axioms in classical logic is trivial. Every sequent whatsoever follows from $\{\Rightarrow_1 p_1, \Rightarrow_1 \neg p_1\}$ in classical logic, including $\Rightarrow_1 A$ for all formulae A. In classical logic, there is no way for $\Rightarrow_1 p_1$ and $\Rightarrow_1 \neg p_1$ to be satisfied at the same valuation; the two inferences are not *jointly satisfiable*. Therefore it is trivially true that A is true in all models in which p_1 and $\neg p_1$ are true, and so any sentence A is a member of the theory

[10] In doing so, we need not take $\Rightarrow_1 p_1$ to be logically valid. We can take $\Rightarrow_1 p_1$ as an axiom simply as a way to restrict to models that satisfy p_1. We can do this because in both M_∞ and classical logic, a model is a counterexample to a formula A iff it is a counterexample to $\Rightarrow_1 A$ iff it is a counterexample to $\Rightarrow_2\Rightarrow_1 A$, and so on.

[11] I take this to be one way of adding axioms to a sequent calculus system, but it is not the only way. For my purposes, it is enough that this is *one* potentially useful way of adding non-logical axioms to the system, which we may want to use for some kinds of axioms.

of p_1 and $\neg p_1$ in classical logic.

However, this is not the case in M_∞. To see why, note that both $\Rightarrow_1 p_1$ and $\Rightarrow_1 \neg p_1$ will be satisfied at any SV model in which at least one Boolean valuation assigns p_1 value 1, and at least one valuation assigns p_1 value 0. Such models need not be trivial in M_∞; any set of two Boolean valuations assigning different values to p_1 will do. $\Rightarrow_1 p_1$ and $\Rightarrow_1 \neg p_1$ are therefore *jointly satisfiable* in M_∞: there are models that satisfy both inferences. This means that the question of which formulae and inferences are satisfied at the models of $\{\Rightarrow_1 p_1, \Rightarrow_1 \neg p_1\}$ is not at all trivial. For some models m and some formulae A, there will be no valuation in m that assigns value 1 to A. For example, there is no Boolean valuation that assigns value 1 to $p_1 \wedge \neg p_1$. As a result, $\Rightarrow_1 p_1 \wedge \neg p_1$ is antivalid in M_∞: there is no model that satisfies $\Rightarrow_1 p_1 \wedge \neg p_1$. This is true even if we restrict our attention to models that satisfy both $\Rightarrow_1 p_1$ and $\Rightarrow_1 \neg p_1$. Therefore $\Rightarrow_1 p_1 \wedge \neg p_1$ is not in the theory of $\Rightarrow_1 p_1$ and $\Rightarrow_1 \neg p_1$ in M_∞ (even though $\{\Rightarrow_1 p_1, \Rightarrow_1 \neg p_1\} \Rightarrow_2 \{\Rightarrow_1 p_1 \wedge \neg p_1\}$ is valid). But it *is* in the theory of those axioms in classical logic; everything is. As such, the same axioms in the same language can have different theories in M_∞ than in classical logic. M_∞ and classical logic are therefore distinct logics.

2.5 Identity Conditions for Logics

I've argued that classical logic and M_∞ are not the same logic, despite the fact that they have the same validities and antivalidities at every inferential level. We might then ask, under what conditions can logics L and L′ rightfully be said to be the *same* logic?

Together with the results of [2] and [36], I take the results here to show that having the same validities and antivalidities is not sufficient to identify two formal systems as the same logic, even if they have the same validities and antivalidities at every inferential level. M_∞ and classical logic have exactly the same validities and antivalidities at every inferential level, and yet they can behave quite differently when given the same set of axioms. But the consequences of a set of axioms are not presentation-dependent features of a logic; the same axioms should not generate different theories depending on how we present the logic. M_∞ and classical logic must therefore be distinct logics, and so it is possible for two distinct logics to have all the same validities and antivalidities at every level of inference.

We must look beyond validities and antivalidities in order to determine whether two formal systems L and L′ are distinct logics, or are simply two presentations of the same logic. In light of the discussion in the previous section, I propose that we at least need to consider the sets of inferences that are jointly satisfiable in a given logic. Even if two logics agree on which inferences have counterexamples and non-counterexamples, the logics can still disagree regarding which *sets* of inferences share a single counterexample and which do not. In classical logic, the 1-inferences $\Rightarrow_1 A$ and $\Rightarrow_1 \neg A$ are not jointly satisfiable. In M_∞, they are. This seems to be precisely the reason that classical logic and M_∞ behave differently given p_1 and $\neg p_1$ as axioms: the axioms are jointly satisfiable in M_∞, but are not jointly satisfiable in classical logic. Having the same sets of jointly satisfiable inferences therefore seems promising as an identity condition to distinguish logics.

In fact if we take sets of jointly satisfiable inferences as an identity condition, then the antivalidities condition can be dropped. It is subsumed under the jointly satisfiable sets condition: if two logics have exactly the same sets of jointly satisfiable inferences, then they must also have the same antivalidities. This is because the α-inference Φ is antivalid iff Φ is not satisfiable iff the singleton set of inferences $\{\Phi\}$ is not jointly satisfiable.[12] However, the validities and jointly satisfiable sets conditions are independent. To see why, consider the

[12] Sets of inferences that are not jointly satisfiable therefore generalize the notion of antivalid inference: they are, in effect, the sets of inferences that are together antivalid. We could apply an analogous generalization to validity, and look at the sets of inferences that do not share a single counterexample. L, the local supervaluationist consequence relation, for example, has exactly the same valid 1-inferences as classical logic, but has different sets of 1-inferences that do not share a counterexample: $\Rightarrow_1 A$ and $\Rightarrow_1 \neg A$ cannot share a counterexample in classical logic, but they can in the local supervaluationist logic. Whether or not this generalization has any philosophically interesting applications remains to be seen.

trivial logic in which every inference is valid. In this logic, every set of inferences is jointly satisfiable. In at least some languages, every set of 1-inferences is jointly satisfiable in the strict-tolerant logic ST. This is because (without logical constants and the like) ST has a model in which every formula gets value $\frac{1}{2}$, and every 1-inference is satisfied at that model. So ST and the trivial logic have exactly the same sets of jointly satisfiable inferences, yet they have different validities: some inferences are invalid in ST, but no inferences are invalid in the trivial logic. The same-validities condition therefore cannot be subsumed under the same-sets-of-jointly-satisfiable-inferences condition. Joint satisfiability is a matter of there being a non-counterexample, and the existence of non-counterexamples cannot by itself tell us whether there are *no* counterexamples.

It is worth noting that the new identity condition offered here is a semantic condition. Unlike validity (and possibly antivalidity), "sets of jointly satisfiable inferences" is an inherently semantic notion, defined in terms of models and satisfaction conditions. Some, especially those who take a purely instrumentalist approach to the models for a logic, may object to this as an identity condition on logics.[13]

However, the identity condition offered here is a semantic condition in part because, like the slice-hierarchy logics that came before it, M_∞ is constructed model-theoretically. As such, its consequence relation is defined semantically. There may be proof theoretic ways to formulate these logics without appealing to any model-theoretic definitions. Once this is done, it will hopefully be clear what the proof-theoretic equivalent of jointly satisfiable inferences might be; it may be some sort of closure operator on sets of non-logical axioms. By introducing this semantic identity condition, I do not mean to suggest that there is no equivalent proof-theoretic condition that could serve the same purpose. There may well be a way to distinguish M_∞ from classical logic without making any appeal to models or satisfaction conditions. The important point is that, whatever that equivalent proof-theoretic

[13]Thanks to an anonymous reviewer for raising this issue.

condition might be, it will have to be go beyond valid and antivalid inferences.

It is also worth noting that sets of jointly satisfiable inferences, as an identity condition on logics, is more fine-grained than the usual properties used to identify logics, like the set of valid inferences or a counterexample relation.[14] One lesson we can learn from comparing M_∞ and classical logic (or comparing ST_∞ and classical logic) is that the valid and antivalid inferences of a logic do not by themselves tell us what consequences a set of axioms will have in the logic. M_∞ and classical logic have exactly the same valid inferences, yet in some cases they will give us different consequences for the same set of axioms in the same language.

If we were to look only at validities, or only at validities and antivalidities, then we would not have enough information to determine what we could or could not prove in the logic from non-logical axioms. So in order to understand how the logic behaves and what is or isn't provable in the logic, we need to look to more fine-grained details of the logic beyond just which inferences are valid. Valid inferences alone are too coarse-grained to tell us what follows from a set of axioms in the logic, and are therefore too coarse-grained to tell us whether or not two formal systems are in fact the same logic. As such, we need to look to more fine-grained distinctions in order to determine whether two formal systems will give us the same consequences for the same set of axioms.

2.6 Paraconsistency and Nontransitive Consequence Relations

2.6.1 Is M_∞ Paraconsistent?

In [36], Scambler argues that the hierarchy logic based on ST introduced in [2] and [28] as a fully classical logic, which I will call ST_∞, is not in fact classical logic. He argues that, unlike

[14]Thanks to an anonymous reviewer for raising this issue.

classical logic, ST_∞ is really a paraconsistent logic: "In the case of $[ST_\infty]$, we have not really gotten rid of paraconsistency: we have merely thoroughly repressed it, so that it does not affect validity at any orders. Nevertheless, it is still present: there are valuations on which $p \wedge \neg p$ comes out valid" [36].

I take Scambler's point here to be something like the following: when we discuss paraconsistency, we often discuss it in terms of the validity or invalidity of the various rules of Explosion, like $A, \neg A \Rightarrow_1 B$ and $A \wedge \neg A \Rightarrow_1 B$. However, in these discussions, we are not interested in the validity of these schema for their own sake. Part of our interest in the validity or invalidity in these schema is that we take them to give us information as to whether or not the logical system can tolerate inconsistency. But ST_∞ and M_∞ can tolerate inconsistencies in their models: both logics have models at which both $\Rightarrow_1 A$ and $\Rightarrow_1 \neg A$ are satisfied. This, Scambler suggests, means that they are really paraconsistent logics.

Normally, a logic is called "paraconsistent" only if some version of Explosion (usually $A, \neg A \Rightarrow_1 B$ or $A \wedge \neg A \Rightarrow_1 B$) is invalid in that logic. But ST_∞ and M_∞ validate every rule of Explosion that classical logic validates, including the metainferential Explosion rule discussed in [3]. As such, ST_∞ and M_∞ are neither strongly nor weakly paraconsistent, in Hyde's terminology [21], [20].[15] So any by the usual definitions of "paraconsistent", ST_∞ and M_∞ simply are not paraconsistent.

However, per Scambler's point, ST_∞ and M_∞ certainly have some paraconsistent-ish features. Both logics have models at which both $\Rightarrow_1 A$ and $\Rightarrow_1 \neg A$ are satisfied.[16] As a result, the logics can tolerate inconsistent axioms in a way that classical logic cannot. In classical logic, theories are closed under Explosion, in the following sense: if $A, \neg A \in T$ for theory T, then $B \in T$. Although Explosion is valid in ST_∞ and M_∞, theories are not closed

[15] Hyde credits this distinction to Arruda [1]. Equivalently, we could say in Ripley's terminology that the two logics are neither *conjunctively* nor collectively paraconsistent [32].

[16] It is worth noting that, although M_∞ has no models at which both A and $\neg A$ get value 0, L_∞ does. L_∞ is therefore in a similar situation: it validates the Law of Excluded Middle, yet has models that satisfy neither $\Rightarrow_1 A$ nor $\Rightarrow_1 \neg A$.

under Explosion in these logics: there are theories containing A and $\neg A$ but not B, for some sentences A and B. So although ST_∞ and M_∞ are not paraconsistent by the usual definitions of paraconsistency, they should be considered at least psuedo-paraconsistent logics.

2.6.2 Scambler's Tortoise Objection to ST_∞ and M_∞

Scambler argues that this paraconsistency (or pseudo-paraconsistency) shows that ST_∞ (and by analogy, M_∞) cannot really be considered a presentation of classical logic. On this point, Scambler and I are in agreement. But Scambler further argues that there is something *wrong* with these logics. He argues that hierarchy logics like ST_∞ are not "closed under their own laws" in an important sense, and that this raises potential problems not just for any proponents of these logics, but for proponents of ST and similar logics as well.

Scambler [36] compares ST_∞ to Lewis Carroll's Tortoise [5]. The Tortoise accepts A, and B, and accepts C := "if A and B are true, Z must be true" but still does not accept Z. The Tortoise continues to accept statements of the form "if A and B and C and... are true, then Z must be true", but the Tortoise still refuses to accept Z.

I take the primary lesson of Carroll's paper to be that accepting or endorsing the *statement* of a rule is very different from actually *obeying* a rule. Asserting that an inference is valid is not the same thing as actually making that inference. Scambler's objection to ST_∞, and by extension to M_∞, is that these logics in effect "accept" classical inferences as valid, without in fact allowing us to *make* those inferences.

Scambler illustrates this issue by introducing a liar constant to the language, but we can make the same point without moving to a new language by looking at inconsistent axioms. ST_∞ and M_∞ both validate the Explosion Rule $A, \neg A \Rightarrow_1 B$, as well as the Metainferential Explosion Rule $\{\Rightarrow_1 A, \Rightarrow_1 \neg A\} \Rightarrow_2 \{\Rightarrow_1 B\}$. However, if we take A and $\neg A$ to be axioms (or $\Rightarrow_1 A$ and $\Rightarrow_1 \neg A$, in our sequent calculus presentation), we see that the theory generated by these axioms is not closed under explosion. M_∞, for example, has models in which $\Rightarrow_1 A$

and $\Rightarrow_1 \neg A$ are satisfied, but $\Rightarrow_1 B$ is not. The theory generated by these axioms in M_∞ therefore does *not* contain every sentence B whatsoever.

Scambler says that ST_∞ is not "closed under its own laws". For our purposes, we might instead say that theories in ST_∞ and M_∞ are not closed under valid inferences. The end result, however, is the same: the valid inferences of these logics do not necessarily correspond to *rules of inference* that we can use when proving theorems from non-logical axioms. Like Carroll's Tortoise, these logics accept the Explosion Rule, but do not allow us to infer *according to* the Explosion Rule.

Scambler [36] suggests that this poses a potential problem for these logics. ST_∞ and M_∞ seem to make precisely the same move that the Tortoise makes: endorsing rules that one does not follow. Defenders of these logics would therefore seem to be endorsing inferences without actually *making* those inferences. But then it is not clear exactly why one would want a logic that validates rules of inference that one cannot use. This certainly does appear to be a problem for these logics. Any defenders of ST_∞ and M_∞ would have to explain what purpose validating rules of inference that we cannot use might have.

2.6.3 The case of ST

Scambler argues that this is not just a problem for hierarchy logics like ST_∞ and M_∞; he argues that it also poses a problem for ST. ST has a nontransitive consequence relation for 1-inferences, but transitive consequence relations for every higher level of inference.[17] As Scambler puts it, "If the problem [with logics like ST_∞ and M_∞] is (as I suggested) that the logic is not closed under its own laws, then why isn't the fact that logics like [ST] also aren't closed under their laws similarly problematic? Don't we have essentially the same structure in each case?" [36].

Scambler's objection, I take it, is this: if nontransitive consequence relations are simply

[17] Recall that by "transitive", I mean that $\vDash_m \Gamma \Rightarrow_1 A$ and $\vDash_m A \Rightarrow_1 \Delta$ imply $\vDash_m \Gamma \Rightarrow \Delta$.

endorsing rules that they don't obey, then this seems like it will be a problem at *any* level. The defenders of ST, who endorse a nontransitive consequence relation at the level of 1-inferences, therefore have to explain why they are in any better position than Carroll's Tortoise, or the hierarchy logics that have nontransitive consequence relations at *every* level.

I agree with Scambler that the full hierarchy logics like ST_∞ and M_∞ seem to face a serious problem that defenders of those logics would have to address. But I do not think that this is *necessarily* a problem for defenders of ST.

In the case of ST, this is potentially a serious problem *if* we want to use the 1-inferences of ST to reason normally. Because ST holds premises to a different standard than it holds conclusions, valid 1-inferences do not preserve any nice properties like truth in a model (i.e. 0-inference satisfaction). In particular, this means that the theory of a set of axioms is not necessarily closed under the valid 1-inferences in ST.[18] The 1-inference $A, \neg A \Rightarrow_1 B$ is valid in ST, yet there are models of ST in which both A and $\neg A$ are satisfied 0-inferences but B is not. So if we try to use the valid 1-inferences of ST as rules of inference applied to axioms, we will end up "proving" sentences that do not actually follow from those axioms in ST.

As I said, this is *potentially* a serious problem for ST. However, this is not a problem if we want to use ST in a different way. For example, Ripley [33] presents a bilateralist interpretation of ST, according to which validity is understood in terms of assertion and denial.[19] On Ripley's bilateralist interpretation of ST, a 1-inference $\Gamma \Rightarrow_1 \Delta$ is valid iff the "position" of asserting all of the γs and denying all of the δs is incoherent. So according to Ripley, "$\Gamma \Rightarrow_1 \Delta$ can now be read as *the claim that* the position [of asserting the γs and denying the δs] is out of bounds." (emphasis mine; notation slightly altered) [33].

[18] Valid 1-inferences in ST *do* have the disjunctive property of either preserving what we might call "tolerant truth" (having value 1 or $\frac{1}{2}$) from left to right *or* preserving "tolerant untruth" (having value 0) from right to left. Unfortunately, this does not suffice to close theories under the valid 1-inferences. This is in part because the validity of an inference does not by itself tell us which property is preserved by that inference. As a result, neither property is preserved in all cases. Thanks to an anonymous reviewer for raising this issue.

[19] For earlier defenses of bilateralism independent of ST, see [35] and [30].

Thus on the bilateralist interpretation of ST, valid 1-inferences need not be understood as representing rules of inference that are safe to use. Rather, they should be understood as claims about what is impermissible to assert and deny. ST is then not a tool for reasoning about *sentences*, but a tool for reasoning about *positions*. On this interpretation, the 1-inferences are the claims *about which* we are making inferences; they do not themselves correspond to rules that we use to make inference.

In some sense, this means that the 1-inferences of ST cannot be used in the way that we usually use inferences. But ST is still a perfectly usable logic for reasoning about positions, because the valid 2-inferences preserve 1-inference validity. Furthermore, the theory generated by any set of 1-inferences that we take as axioms will be closed under the (locally) valid 2-inferences of ST. If we take a set of 1-inferences as axioms, we can therefore use (locally) valid 2-inferences as rules of inference.

This interpretation of ST therefore avoids Scambler's objection, because it does not endorse rules that it refuses to obey. It obeys the 2-inference rules that it endorses, and it does not consider valid 1-inferences to be rules at all.[20]

However, this same approach will not work for ST_∞ or M_∞. It is crucial to the bilateralist account of ST that we can understand metainferences in the usual way: as formal representations of rules of inference that we can use. Without that, it's not clear how we could use the logic *as a logic*. But ST_∞ and M_∞ have nontransitive consequence relations at *every level of inference*. As a result, there is no level of inference at which the valid inferences can be understood as rules of inference: there is no ordinal α at which theories are closed under all valid α-inferences. The bilateralist interpretation of ST reinterprets valid inferences at one level as claims instead of rules, but we can still use the valid inferences of the next level as rules of inference. In ST_∞ and M_∞, *every* level has to be reinterpreted. This leaves no

[20]This is not to say that the bilateralist interpretation is free from objections; only that it is free from this particular objection.

level at which valid inferences can be understood as rules of inference that it is safe to use. It's therefore not clear how we are to use these logics, if we are to use them.

There may be uses for such logics; I leave open the question of whether or not formal systems like M_∞ can serve useful purposes. There are many ways to use a formal construction. For example, in this paper I have used M_∞ as an example in an argument for certain claims about identity criteria for logics. But in doing so, I wasn't really using M_∞ *as a logic*. To use a logic as a logic, "from the inside" so to speak, we need to be able to use the formal construction as a tool for making inferences. I take Scambler's objection to hierarchy logics like ST_∞ and M_∞ to be that, due to the non-transitive nature of the logics, they cannot really be used as logics in this sense. In that, I agree.

That is not to say that these logics cannot be useful; there may be many purposes for which we could use them. But they cannot be used in what we might call the "usual way". There are important uses of a logic for which hierarchy logics simply will not work. What other purposes these logics might serve depends in part on how these logics can be usefully interpreted. At the moment, it is not clear how this is to be done. And this does indeed present a problem for any defenders of these logics.

2.7 Conclusion

I've argued that a logic cannot be identified by its valid and antivalid inferences, even at every inferential level. At a minimum, we suggest that we must also look to which sets of inferences are jointly satisfiable. Ultimately, we need to look to the rules of inference that the logic allows us to use. Two logics having exactly the same validities and antivalidities does not suffice to guarantee that the two logics obey the same rules of inference, or that they will allow us to prove the same consequences from the same set of axioms.

I take it that when we attempt to characterize logics and logical properties by the validity or invalidity of inferences, we often do so because we make certain assumptions about what those inferences represent. The recent development of mixed-condition consequence relations has demonstrated that these assumptions can be broken. In particular, the valid inferences of a logic can come apart from the rules of inference that the logic allows us to use. This means that instead of looking only at the inferences of a logic, we should be looking directly at the properties and rules of inference that the logic allows. Although inferences *can* represent these properties and rules in some settings, I take the results here and in [2] and [36] to show that inferences do not always do so.[21]

[21]Earlier versions of this paper were presented at the Buenos Aires Logic Group WIP Seminar, and at a workshop on substructural logics and hierarchies thereof hosted by the CUNY Logic and Metaphysics Workshop and the Saul Kripke Center. I am grateful to the audiences of these talks for their feedback. I am indebted to Eduardo Barrio, Paul Egré, Federico Pailos, Graham Priest, David Ripley, and Chris Scambler for many helpful discussions of these ideas and for all of their feedback on earlier drafts of the paper. Many thanks also to the two anonymous referees, whose helpful comments improved the paper significantly.

Bibliography

[1] Arruda, A. (1989) "Aspects of the Historical Development of Paraconsistent Logic" in *Paraconsistent Logic: Essays on the Inconsistent*, ed. Priest, Routley and Norma. Munchen: Philosophia.

[2] Barrio, E., Pailos, F., Szmuc, D. (2019) "A Hierarchy of Classical and Paraconsistent Logics" *Journal of Philosophical Logic*, 49: 93-120.

[3] Barrio, E., Pailos, F., Szmuc, D. (2018) "What is a Paraconsistent Logic?", in *Contradictions, from Consistency to Inconsistency*, ed. J. Malinowski and W. Carnielli. Springer Verlag.

[4] Boghossian, P. (2014) "What is Inference?" *Philosophical Studies*, 169(1): 1-18.

[5] Carroll, L. (1895) "What the Tortoise Said to Achilles" *Mind*, 4(14), 278-280.

[6] Cobreros, P. (2013) "Vagueness: Subvaluationism" *Philosophy Compass* 8(5): 472-485.

[7] Cobreros, P. (2011) "Paraconsistent Vagueness: A Positive Argument" *Synthese*, 183(2): 211-227.

[8] Cobreros, P., Egré, P., Ripley, D., van Rooij, R. (2012) "Tolerant, classical, strict." *Journal of Philosophical Logic*, 41, 347-385.

[9] Cobreros, P., Egré, P., Ripley, D., van Rooij, R. (2013) "Reaching Transparent Truth." *Mind*, 122(488), 841-866.

[10] Cobreros, P., Egré, P., Ripley, D., van Rooij, R. (2012) "Tolerance and mixed consequence in the S'valuationist setting." *Studia Logica*, 100(4), 855–877.

[11] Došen, K. (1985a) "Sequent-Systems for Modal Logic" *Journal of Symbolic Logic*, 50: 149-159.

[12] Došen, K. (1986) "Higher-Level Sequent Systems for Intuitionistic Modal Logic" *Publications de l'Institut Mathématique*, 39(53): 3-12.

[13] Došen, K. (1988) "Sequent Systems and Groupoid Models I" *Studia Logica*, 47: 353-389.

[14] Došen, K. (1992) "Modal Logic as Metalogic" *Journal of Logic, Language, and Information* 1: 173-201.

[15] Fine, K. (1975) "Vagueness, Truth and Logic" *Synthese*, 30: 265-300.

[16] Fitting, Melvin (2021a) "The Strict/Tolerant Family Continued: Quantifiers and Modalities" *The Australasian Journal of Logic*, 18(6): 616-644.

[17] Fitting, Melvin (2021b) "A Family of Strict/Tolerant Logics" *Journal of Philosophical Logic*, 50: 363–394.

[18] Fitting, Melvin (2021c) "The Strict/Tolerant Idea and Bilattices" in *For Arnon Avron*, ed. Ofer Arieli et al., Outstanding Contributions to Logic.

[19] Humberstone, L. (1996) "Valuational Semantics of Rule Derivability" *Journal of Philosophical Logic*, 25(5): 451-461.

[20] Hyde, D. (2008) *Vagueness, Logic and Ontology*. Aldershot: Ashgate.

[21] Hyde, D. (1997) "From Heaps and Gaps to Heaps of Gluts" *Mind*, 106(424): 641-660.

[22] Hyde, D. and Colyvan, M. (2008) "Paraconsistent Vagueness: Why not?" *Australasian Journal of Logic*, 6: 107-121.

[23] Keefe, R. (2008) "Vagueness: Supervaluationism" *Philosophy Compass* 3: 315-324.

[24] Keefe, R. (2000) *Theories of Vagueness*. Cambridge: Cambridge University Press.

[25] Kripke, S. (1975) "Outline of a Theory of Truth" *Journal of Philosophy*, 72: 690-716.

[26] Maudlin, T. (2004) *Truth and Paradox*. Oxford University Press.

[27] McGee, V. and McLaughlin, B. (1995) "Distinctions without a Difference" *Southern Journal of Philosophy*, 33: 203-251.

[28] Pailos, F. (2019) "A fully classical truth theory characterized by substructural means" *The Review of Symbolic Logic*.

[29] Priest, G. "Paraconsistent Logic", in *Handbook of Philosophical Logic*, second edition, Volume 6, ed. D. Gabbay and F. Guenther, pp. 287-393.

[30] Restall, G. (2005) "Multiple Conclusions", in *Logic, Methodology, and Philosophy of Science: Proceedings of the Twelfth International Congress*, ed. P. Hajek, L. Valdes-Villanueva, and D. Westerståhl. London: Kings College Publications: 189–205.

[31] Ripley, D. (2018) "On the 'Transitivity' of Consequence Relations" *Journal of Logic and Computation*, 28(2):433-450.

[32] Ripley, D. (2015) "Paraconsistent Logic" *Journal of Philosophical Logic*, 44: 771-780.

[33] Ripley, D. (2013) "Paradoxes and Failures of Cut" *Australasian Journal of Philosophy* 91 (1): 139-164.

[34] Ripley, D. (2012) "Conservatively Extending Classical Logic with Transparent Truth" *Review of Symbolic Logic*, 5 (2): 354-378.

[35] Rumfitt, I. (2000) ""Yes" and "No"" *Mind*, 104(436): 781–823.

[36] Scambler, C. (2020) "Classical Logic and the Strict Tolerant Hierarchy" *Journal of Philosophical Logic*, 49: 351-370.

[37] Scambler, C. (forthcoming) "Transfinite Meta-inferences" *Journal of Philosophical Logic*, https://doi.org/10.1007/s10992-020-09548-7.

[38] Van Fraassen, B. C. "Singular Terms, Truth-Value Gaps, and Free Logic" *Journal of Philosophy*, 63: 481-495.

[39] Varzi, A. (2007) "Supervaluationism and its Logics" *Mind*, 116: 633-676.

[40] Varzi, A. (1997) "Inconsistency Without Contradiction" *Notre Dame Journal of Formal Logic*, 38: 621-640.

[41] Williamson, T. (1994) *Vagueness*. London: Routledge.

Chapter 3

Infinite-Premise Paradoxes and Contraction

3.1 Introduction

In recent years, substructural logics have been presented as potential solutions to the semantic paradoxes. These proposed solutions avoid the paradoxes by giving up one of the standard structural rules, either Cut or Contraction. Nontransitive approaches give up Cut in order to tolerate the inconsistencies that arise in a naive theory of truth without triviality.[1] Noncontractive logics give up Contraction in order to avoid inconsistencies in the first place. Both sorts of substructural systems can provide a naive theory of truth without triviality. These substructural approaches aim to provide a uniform solution to the semantic paradoxes, in which every paradoxical argument is blocked at the exact same step: either an application of Cut, or an application of Contraction.

If these substructural approaches are going to provide a uniform solution to semantic paradoxes, then these substructural logics need to provide logics that we can use in the way that we usually use logics. One important use of a logic is to discover the *consequences* of a set of axioms, premises, or assumptions. For example, in mathematics we are often interested in the consequences of ZFC. In philosophy, we are interested in the consequences of particular theories of truth, or knowledge, or ethics. In these cases, we want to know what our assumptions or axioms make *true*, or what we are obligated to accept *given* that we accept those assumptions. Crucially, the collections of axioms in which we are interested are often infinite: we will never use *all* of the axioms in any single proof, but we want to use our logic to discover the consequences of this infinite collection of axioms.

I have argued elsewhere, in "Supervaluations and the Strict-Tolerant Hierarchy", that at least some nontransitive approaches cannot be used for this purpose. Specifically, I have argued that any logic that goes nontransitive "all the way" through the infinite hierarchy of metainferences cannot be used to discover the consequences of a set of nonlogical axioms.

[1]For example: the nontransitive logic STT validates $\vdash L \wedge \neg L$—where L is the liar sentence— and also validates $L \wedge \neg L \vdash B$. But without Cut, this does not imply $\vdash B$, and thus triviality is avoided.

In this paper, I argue that noncontractive approaches fare no better. In section 3.2, I discuss noncontractive logics, and some examples of the noncontractive approach. In section 3.3, I present paradoxes involving an infinitary conjunction, which cannot be avoided by dropping the structural rule of Contraction. In section 3.4, I present a version of the validity curry paradox that involves countably infinitely many premises. In section 3.5, I present a version of the validity curry paradox for what Mares and Paoli [10] call "external validity". In section 3.6, consider a possible response on behalf of the the noncontractive approach, and argue that it does not avoid paradox. In section 3.7, I argue that these failures are part of a general problem that noncontractive approaches have with *closure*. I consider a recent proposal that aims to solve this problem, and argue that it fails. In section 3.8, I conclude that the noncontractive approach does not provide a uniform solution to semantic paradoxes. There are contexts in which we need a logic to discover the consequences of nonlogical axioms, and for those contexts the noncontractive approach tells us nothing; it merely changes the subject.

3.2 Noncontractive Logics

A noncontractive logic is a logic in which the structural rules of Contraction, WL and WR:

$$\frac{\Gamma, A, A \vdash \Delta}{\Gamma, A \vdash \Delta}\text{WL}$$

$$\frac{\Gamma \vdash B, B, \Delta}{\Gamma \vdash B, \Delta}\text{WR}$$

This rule is used, explicitly or implicitly, in the derivation of virtually every semantic paradox: the liar, Curry, validity curry, and many more all involve some essential use of Contraction. By dropping contraction, philosophers and logicians have aimed to provide a single uniform solution to the semantic paradoxes: seemingly any semantic paradox can be solved by dropping Contraction.[2]

Consider the following derivations of the Liar, Curry, and Validity Curry paradoxes:

Liar Paradox

Let λ be a sentence such that the term $\langle\lambda\rangle$ is the name for the sentence $T(\langle\lambda\rangle)$. The paradox can be derived as follows:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\lambda \vdash \lambda}{\lambda \vdash T(\langle\lambda\rangle)}\text{T-intro}}{\lambda, \neg T(\langle\lambda\rangle) \vdash}\neg\text{L}}{\neg T(\langle\lambda\rangle), \neg T(\langle\lambda\rangle) \vdash}\text{def. of }\lambda}{\neg T(\langle\lambda\rangle) \vdash}\text{Contraction} \qquad \dfrac{\dfrac{\dfrac{\dfrac{\lambda \vdash \lambda}{T(\langle\lambda\rangle) \vdash \lambda}\text{T-intro}}{\vdash \lambda, \neg T(\langle\lambda\rangle)}\neg\text{R}}{\vdash \neg T(\langle\lambda\rangle), \neg T(\langle\lambda\rangle)}\text{def. of }\lambda}{\vdash \neg T(\langle\lambda\rangle)}\text{Contraction}}{\vdash}\text{Cut}$$

Curry's Paradox

Let κ be a sentence such that the term $\langle\kappa\rangle$ is the name for the sentence $T\langle\kappa\rangle \to \bot$. The paradox can be derived in two parts:

Part A is a derivation of $T\langle\kappa\rangle \vdash \bot$:

$$\dfrac{\dfrac{\dfrac{\dfrac{T(\langle\kappa\rangle) \vdash T(\langle\kappa\rangle) \qquad \bot \vdash \bot}{T\langle\kappa\rangle, T\langle\kappa\rangle \to \bot \vdash \bot}\to\text{L}}{T\langle\kappa\rangle, \kappa \vdash \bot}\text{def. of }\kappa}{T\langle\kappa\rangle, T\langle\kappa\rangle \vdash \bot}\text{T-intro}}{T\langle\kappa\rangle \vdash \bot}\text{Contraction}$$

Part B uses two occurrences of Part A to derive $\vdash \bot$:

$$\dfrac{\dfrac{\dfrac{\dfrac{\text{[Part A]}}{T\langle\kappa\rangle \vdash \bot}}{\vdash T\langle\kappa\rangle \to \bot}\to\text{R}}{\vdash \kappa}\text{def. of }\kappa \qquad \dfrac{\text{[Part A]}}{T\langle\kappa\rangle \vdash \bot}}{\vdash \bot}\text{Cut}$$

[2]See, for example, [14], [26], [27], [19], [20], [4], [10], [25], and [5].

Validity Curry Paradox

Let $Val(x, y)$ be a predicate representing validity, which is governed by the following two rules:

Validity Proof:

$$\frac{\Gamma \vdash \Delta}{\vdash Val(\langle\Gamma\rangle, \langle\Delta\rangle)}\ \text{VP}$$

Validity Detachment:

$$\frac{}{\Gamma, Val(\langle\Gamma\rangle, \langle\Delta\rangle) \vdash \Delta}\ \text{VD}$$

Let π be a sentence such that $\langle\pi\rangle$ is a name for the sentence saying that the inference from π to $\bot$ is valid: $Val(\langle\pi\rangle, \langle\bot\rangle)$. Like the original Curry paradox, $\vdash \bot$ can be derived in two parts:[3]

Part A is a derivation of $\pi \vdash \bot$, using only (VD) and Contraction:

$$\dfrac{\dfrac{\dfrac{}{\pi, Val(\langle\pi\rangle, \langle\bot\rangle)\ \vdash\ \bot}\ \text{VD}}{\pi, \pi\ \vdash\ \bot}\ \text{def. of } \pi}{\pi\ \vdash\ \bot}\ \text{Contraction}$$

Part B uses Part A, along with (VP) and Cut, to derive $\vdash \bot$:

$$\dfrac{\dfrac{\dfrac{\dfrac{[\text{Part A}]}{\pi\ \vdash\ \bot}}{\vdash\ Val(\langle\pi\rangle, \langle\bot\rangle)}\ \text{VP}}{\vdash\ \pi}\ \text{def. of } \pi \qquad \dfrac{[\text{Part A}]}{\pi\ \vdash\ \bot}}{\vdash\ \bot}\ \text{Cut}$$

All three of these paradoxes, and many others, require Contraction. Dropping Contraction blocks these derivations, and blocks each derivation at the exact same step. And in fact, several noncontractive logics have been proven to be consistent and nontrivial; there is no

[3] This presentation of the validity curry paradox is based on that of [4].

derivation of $\vdash \bot$.[4] Noncontractive approaches therefore appear to provide a *uniform* solution to semantic paradoxes: they solve the semantics paradoxes, and solve each paradox in the exact same way. Every paradox is given the same diagnosis: an illicit use of Contraction. This is particularly useful in so-called "structural paradoxes" like the validity curry.[5] These paradoxes do not involve logical operators like negations or conditionals, and so cannot be solved by the usual methods employed by paracomplete and paraconsistent logics.

3.2.1 Rejecting Contraction

In order for Contraction to fail, it is important that the premises and conclusions be collected in *multisets*, rather than sets. Sets automatically contract: the set $\{A, A\}$ *just is* the set $\{A\}$. But multisets can have multiple occurrences of the same member: the multiset $[A, A]$ is distinct from the multiset $[A]$.

In sequents with multiple multisets separated by a comma, such as $\Gamma, \Sigma \vdash \Delta$, the comma is understood as multiset *sum*, or $\boxplus$. Γ, Δ is the sum $\Gamma \boxplus \Delta$, which is the multiset that contains every formula that occurs in either Γ or Δ, the number of occurrences of that formula is the *sum* of the number of occurrences in Γ plus the number of occurrences in Δ.[6]

One consequence of this failure is that different ways of formulating the operational rules governing connectives, which are equivalent in the presence of Contraction and Weakening, are no longer equivalent. Different choices of rules will result in different versions of the logical connectives. For example, we might introduce the conjunction $\wedge$ using the following "multiplicative" rules:

$$\frac{\Gamma, A, B \vdash \Delta}{\Gamma, A \wedge B \vdash \Delta}\, m{\wedge}\mathrm{L}$$

$$\frac{\Gamma \vdash A, \Delta \qquad \Gamma' \vdash B, \Delta'}{\Gamma, \Gamma' \vdash A \wedge B, \Delta, \Delta'}\, m{\wedge}\mathrm{R}$$

[4] See, for example, [14] and [26].
[5] See [10] and [4].
[6] See [16] for a very helpful discussion of these issues.

Or alternatively, we might introduce conjunction using the following "additive" rules:

$$\frac{\Gamma, A \vdash \Delta}{\Gamma, A \wedge B \vdash \Delta}\ a{\wedge}\mathrm{L} \qquad\qquad \frac{\Gamma, B \vdash \Delta}{\Gamma, A \wedge B \vdash \Delta}\ a{\wedge}\mathrm{L}$$

$$\frac{\Gamma \vdash A, \Delta \qquad \Gamma \vdash B, \Delta}{\Gamma \vdash A \wedge B \Delta}\ a{\wedge}\mathrm{R}$$

Without Contraction (and Weakening), these rules are not equivalent. For example, suppose we have derived:

$$\Gamma, A, B \vdash \Delta$$

If we only have the additive conjunction rules, then this does not get us:

$$\Gamma, A \wedge B \vdash \Delta$$

Rather, we can only derive:

$$\Gamma, A \wedge B, A \wedge B \vdash \Delta$$

by applying the $a{\wedge}$L rule twice; once to A, and once to B. Thus one occurrence of the additive conjunction of A and B is in some ways logically *weaker* than A and B taken as separate premises. The fact that we have $A, B \vdash \Delta$ does *not* guarantee that we have $A \wedge B \vdash \Delta$ if $\wedge$ obeys the additive conjunction rules.

One consequence of this additive/multiplicative split is that not every classical inference can be recovered by either multiplicative connectives or additive connectives alone. For example, if $\wedge$ is the *additive* version of the conjunction, then although we have one version of Explosion:

$$A, \neg A \vdash B$$

for all A and B, we do *not* have the following version:

$$A \wedge \neg A \vdash B$$

for all A and B. We only have this weaker version:

$$A \wedge \neg A, A \wedge \neg A \vdash B$$

However, if $\wedge$ is the *multiplicative* conjunction, then we do have the usual conjunctive Explosion rule:

$$A \wedge \neg A \vdash B$$

It follows directly from

$$A, \neg A \vdash B$$

by one application of $m\wedge$L. However, the inference

$$A \vdash A \wedge A$$

is not valid for the multiplicative conjunction. You need two occurrences of A to get the multiplicative conjunction $A \wedge A$.

Despite this, there are ways to combine these rules to recover all classically valid inferences. For example, [17] provides a noncontractive system (which is also nontransitive) that has mixed rules (multiplicative on the left and additive on the right) and validates all of the classically valid inferences. In [10], Mares and Paoli defend a system of linear logic that contains all of the multiplicative and additive rules, giving us two versions of each connective: one additive and one multiplicative. They argue that classically valid inferences are all ambiguous between the two sorts of connectives, and show that every classically valid inference has a disambiguation that is valid in their formal system.

This division of connectives into additive and multiplicative versions is important, as it prevents derivations of the paradoxes that don't *directly* involve Contraction. For example, consider the following derivation:

$$\dfrac{\dfrac{\dfrac{\lambda \vdash \lambda}{\lambda, \neg\lambda \vdash}\neg L}{\lambda \wedge \neg\lambda \vdash} m\wedge L \qquad \dfrac{\lambda \vdash \lambda \qquad \lambda \vdash \neg\lambda}{\lambda \vdash \lambda \wedge \neg\lambda} a\wedge R}{\dfrac{\lambda \vdash}{\vdash \neg\lambda}\neg R} \text{Cut}$$

This would make the negation of the liar sentence a theorem; a quick application of the truth rules would get us a contradiction, and triviality. This derivation conflates multiplicative conjunction with additive conjunction, and thus allows us to sneak in a form of Contraction. We can see this in the following derivation:

$$\dfrac{\dfrac{\Gamma, A, A \vdash \Delta}{\Gamma, A \wedge A \vdash \Delta} m\wedge L \qquad \dfrac{A \vdash A \qquad A \vdash A}{A \vdash A \wedge A} a\wedge R}{\Gamma, A \vdash \Delta} \text{Cut}$$

In effect, conflating the two sorts of connective rules would allow us to sneak a hidden form of contraction into proofs. In the example shown here, the *additive* conjunction of A and A is equivalent to once occurrence of A, but the multiplicative conjunction of A and A is equivalent to *two* occurrences of A. Conflating the two sorts of conjunction is thus really a form of Contraction. It is therefore vital that the two versions of the connective rules be kept separate in a noncontractive system.

Different noncontractive logics will validate different combinations of these connective rules. The noncontractive system presented in [26], for example, uses exclusively multiplicative connective rules for the logical connectives and the quantifiers. The system defended in [21] uses exclusively additive connective rules. The system presented in [17] uses multiplicative rules on the left side of the $\vdash$, and additive rules on the right side (but avoids the above derivation by *also* rejecting Cut). And the system defended in [10] uses both sets of rules, and therefore has two conjunction connectives, two disjunction connectives, etc.

Not all connectives will have an additive/multiplicative split. The rules for negation, for example, have no such division:

$$\neg\text{L}\,\frac{\Gamma \Rightarrow A\Delta}{\Gamma\neg A \Rightarrow \Delta}$$

$$\neg\text{R}\,\frac{\Gamma, A \Rightarrow \Delta}{\Gamma \Rightarrow \neg A\Delta}$$

The case of the quantifiers is unique because, like negation, there is only one traditional set of rules for the quantifier. Yet the usual quantifier rules are distinctly additive, and so are not exactly unsplit in the way that negation is. For example, consider the left-side rule for the universal quantifier:

$$\frac{\Gamma, A(t) \vdash \Delta}{\Gamma, \forall x A(x) \vdash \Delta}\,\forall\text{L}$$

If we think of universal quantification as analogous to a very large conjunction, we can see that this rule is much more like the *additive* left-side conjunction rule than the *multiplicative* left-side conjunction rule. The other quantifier rules are similarly additive in nature. But there is no consensus as to how multiplicative quantifiers should work. In [12], Paoli argues that there should be both multiplicative and additive quantifiers, and makes some suggestions for how multiplicative quantifiers ought to work. In [26], Zardini introduces multiplicative rules for the quantifiers along the lines suggested by Paoli. These quantifiers are unique, in part because the introduction rule for the universal quantifier has infinitely many premises. For example, here are the rules for the universal quantifier (notation adjusted):

$$\frac{\Gamma, A_{t_0/x}, A_{t_1/x}, A_{t_2/x}, A_{t_3/x}, \dots \vdash \Delta}{\Gamma, \forall x A \vdash \Delta}\,\forall\text{L}$$

$$\frac{\Gamma_0 \vdash \Delta_0, A_{t_0/x} \qquad \Gamma_1 \vdash \Delta_1, A_{t_1/x} \qquad \Gamma_2 \vdash \Delta_2, A_{t_2/x} \qquad \Gamma_3 \vdash \Delta_3, A_{t_3/x} \qquad \cdots}{\bigsqcup_{0\leqslant i<\omega}(\Gamma_i) \vdash \bigsqcup_{0\leqslant i<\omega}(\Delta_i), \forall x A}\,\forall\text{R}$$

In Zardini's system, there must be a term t_i for every object in the domain. So if we have $A_{t_i/x}$ for all $i \leqslant \omega$, then we have that every object in the domain is in the extension of A. Thus we can see that these multiplicative quantifier rules are analogous to the multiplicative conjunction rules, just as the usual universal quantifier rules are analogous to the additive conjunction rules. For example in the left side rule, we must already have sentences saying that every object is in the extension of A before we can introduce the universal quantification that effectively "conjoins" those sentences, just as we must already have both conjuncts before we can conjoin them with the multiplicative conjunction:

$$\frac{\Gamma, A, B \vdash \Delta}{\Gamma, A \wedge B \vdash \Delta}\, m\wedge\text{L}$$

Note that these multiplicative quantifier rules require both inferences and metainferences that have infinitely many premises. In the ∀L rule, we have infinitely many premises on the left side of the turnstile. In the ∀R rule, we have infinitely many premises in the metainferential rule itself.[7]

In the following sections, I will present a series of paradoxes do not use the structural rule of Contraction, and thus pose problems for noncontractive approaches given a sufficiently expressive vocabulary. The paradoxes of section 3.3 require that the logic contain some a multiplicative connectives, but the paradoxes of sections 3.4 and 3.5 do not involve logical connectives and apply to any noncontractive logic that tolerates countably infinite collections of premises, regardless of whether the connectives are additive or multiplicative.

3.3 Infinite Conjunction Paradoxes

Let $\bigwedge$ be a multiplicative countable conjunction connective, which we can think of as conjoining all of the members of a (possibly infinite) countable multiset. We will write $\bigwedge\Gamma$ to refer to the conjunction of all of the members of γ.

$\bigwedge$ obeys the following rules:

[7]However, Da Ré and Rosenblatt raise several problems for these quanitifers in [7].Da Ré and Rosenblatt show that Zardini's logical system, when given a transparent truth predicate and a small amount of arithmetic, leads to triviality. They show that their argument will not work for a noncontractive system that contains only the usual "additive" quantifiers, which suggests that the problem is indeed Zardini's quantifiers.

$$\frac{\Gamma \vdash \Delta}{\bigwedge \Gamma \vdash \Delta} \bigwedge\text{L}$$

$$\frac{\Gamma_1 \vdash A_1, \Delta_1 \qquad \Gamma_2 \vdash A_2, \Delta_2 \qquad ... \qquad \Gamma_n \vdash A_n, \Delta_n \qquad ...}{\Gamma_1, \Gamma_2, ...\Gamma_n... \vdash \bigwedge[A_i], \Delta_1, \Delta_2, ...\Delta_n...} \bigwedge\text{R}$$

Given these rules, we can prove that $\bigwedge$ obeys the following introduction rule (at least in the presence of Cut):

$$\frac{\bigwedge \Gamma, A \vdash \Delta}{\bigwedge(\Gamma \boxplus [A]) \vdash \Delta} \bigwedge\text{-intro}$$

For all $\gamma \in \Gamma$, we have:

$$\gamma \vdash \gamma$$

Thus:

$$\frac{\gamma \vdash \gamma \qquad ...}{\Gamma \vdash \bigwedge \Gamma}$$

Thus by Cut, we have:

$$\cfrac{\cfrac{\bigwedge \Gamma, A \vdash \Delta \qquad \Gamma \vdash \bigwedge \Gamma}{\Gamma, A \vdash \Delta}}{\bigwedge(\Gamma \boxplus [A]) \vdash \Delta}$$

In particular, this means that we can have an countably infinite conjunction of occurrences of a sentence A. Importantly, this builds in a sort of contraction to the system. Let $[A]_\omega$ be the multiset consisting of countably infinitely many occurrences of A, and nothing else. If we have $[A]_\omega, A \vdash \Delta$, this is the same thing as $[A]_\omega \vdash \Delta$. This is because, as mentioned earlier, the comma in noncontractive multiset sequent calculi must be understood as multiset *sum*, or $\boxplus$: A formula appears in Γ, Δ the sum of the number of times that it appears in Γ and the number of times it appears in Δ. But in cases where a multiset has infinitely many occurrences of a formula, we will get a de facto Contraction when we add another occurrence; the sum of $\aleph_0$ and 1 is still $\aleph_0$.

Using this infinite conjunction, we can derive paradoxes analogous to the normal liar and curry paradoxes.

An Infinitary Liar Paradox

Let L be a sentence in $\mathcal{L}$ such that $\langle L\rangle$ is the name for the sentence $\bigwedge[\neg T\langle L\rangle]_\omega$. We can divide the paradoxical derivation into four parts.

Part A derives $L \vdash$ using the derived $\bigwedge$ introduction rule, the truth rules, and the rules for negation:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{L \vdash L}{L \vdash T\langle L\rangle}\text{T-intro}}{\bigwedge[\neg T\langle L\rangle]_\omega \vdash T\langle L\rangle}\text{def. of } L}{\bigwedge[\neg T\langle L\rangle]_\omega, \neg T\langle L\rangle \vdash}\neg\text{L}}{\bigwedge([\neg T\langle L\rangle]_\omega \boxplus \neg T\langle L\rangle) \vdash}\bigwedge\text{-intro}}{\bigwedge[\neg T\langle L\rangle]_\omega \vdash}{\scriptstyle ([\neg T\langle L\rangle]_\omega \boxplus \neg T\langle L\rangle) = [\neg T\langle L\rangle]_\omega}}{L \vdash}\text{def. of } L$$

Part B derives $\vdash \neg T\langle L\rangle$ using Part A, truth rules, and the rules for negation:

$$\dfrac{\dfrac{L \vdash}{T\langle L\rangle \vdash}\text{T-intro}}{\vdash \neg T\langle L\rangle}\neg\text{R}$$

Part C derives $\vdash L$ using Part B and $\bigwedge$R:

$$\dfrac{\dfrac{\vdash \neg T\langle L\rangle \qquad \vdash \neg T\langle L\rangle \qquad \ldots}{\vdash \bigwedge[\neg T\langle L\rangle]_\omega}\bigwedge\text{R}}{\vdash L}\text{def. of } L$$

Part D derives the empty sequent $\vdash$ using Parts A and C, and Cut:

$$\dfrac{L \vdash \qquad \vdash L}{\vdash}\text{Cut}$$

This version of the liar paradox does not involve Contraction. The only contraction-like step is the move from $\bigwedge([\neg T\langle L\rangle]_\omega \boxplus \neg T\langle L\rangle) \vdash$ to $\bigwedge[\neg T\langle L\rangle]_\omega \vdash$ in Part A. But this is not Contraction; this is just a fact about countably infinite multisets: the sum of a multiset containing $\aleph_0$-many occurrences of A with a multiset containing one occurrence of A is a multiset containing $\aleph_0$ occurrences of A. Giving up Contraction does nothing to avoid this.

An Infinitary Curry Paradox

Let K be a sentence in $\mathcal{L}$ such that the term $\langle K \rangle$ is the name for the sentence $\bigwedge[T\langle K\rangle]_\omega \multimap \bot$. We can again derive the paradox in four parts.

Part A uses the derived $\bigwedge$ introduction rule, the truth rules, and the multiplicative conditional rules to derive $\bigwedge[T\langle K\rangle]_\omega \vdash \bot$:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\bigwedge[T\langle K\rangle]_\omega \vdash \bigwedge[T\langle K\rangle]_\omega \qquad \bot \vdash \bot}{\bigwedge[T\langle K\rangle]_\omega, \bigwedge[T\langle K\rangle]_\omega \multimap \bot \vdash \bot}\multimap\mathrm{L}}{\bigwedge[T\langle K\rangle]_\omega, K \vdash \bot}\text{def. of } K}{\bigwedge[T\langle K\rangle]_\omega, T\langle K\rangle \vdash \bot}\text{T-intro}}{\bigwedge[T\langle K\rangle]_\omega \boxplus [T\langle K\rangle] \vdash \bot}\bigwedge\text{-intro}}{\bigwedge[T\langle K\rangle]_\omega \vdash \bot}([\neg T\langle K\rangle]_\omega \boxplus \neg T\langle K\rangle) = [\neg T\langle K\rangle]_\omega$$

Part B uses Part A, the truth rules and the multiplicative conditional rules to derive $\vdash T\langle K\rangle$:

$$\dfrac{\dfrac{\dfrac{\bigwedge[T\langle K\rangle]_\omega \vdash \bot}{\vdash \bigwedge[T\langle K\rangle]_\omega \multimap \bot}\multimap\mathrm{R}}{\vdash K}\text{def. of } K}{\vdash T\langle K\rangle}\text{T-intro}$$

Part C uses Part B and $\bigwedge$R to derive $\vdash \bigwedge[T\langle K\rangle]_\omega$:

$$\dfrac{\vdash T\langle K\rangle \qquad \vdash T\langle K\rangle \qquad \ldots}{\vdash \bigwedge[T\langle K\rangle]_\omega}\bigwedge\mathrm{R}$$

Part D uses Parts A and C and Cut to derive $\vdash \bot$:

$$\dfrac{\vdash \bigwedge_\omega T\langle K\rangle \qquad \bigwedge_\omega T\langle K\rangle \vdash \bot}{\vdash \bot}\text{Cut}$$

Like the infinitary liar paradox, this version of the Curry paradox does not rely on Contraction; it only uses basic facts about the sum operator $\boxplus$ for infinite multisets.

Both of these paradoxes use a countable conjunction connective in order to conjoin infinitely many premises into a single sentence, and the infinitary curry paradox involves a multiplicative conditional. But we do not need such a conjunction or such a conditional in order to produce contraction-free paradoxes: we can get infinitary validity curry paradoxes in any noncontractive logic as long as the language contains a validity predicate and names for (possibly infinite) multisets.

3.4 An Infinitary Validity Curry Paradox

[10] distinguishes between *internal* and *external* validity. Roughly, B is an internal consequence of A iff $A \vdash B$, and B is an external consequence of A iff

$$\frac{\vdash A}{\vdash B}$$

In this section, I'll introduce infinite-premise versions of the validity curry paradox for both internal and external validity.

Let $\mathcal{L}$ be a language that contain names not only for formulas, but for countable multisets of formulas. If we allow $\mathcal{L}$ to have names for *all* multisets of formulas, then due to Cantor's Theorem we will have strictly more terms than formulas, which is impossible. To avoid this, while still allowing ourselves to talk about multisets in the object language, we will restrict our set of multiset names to include names only for multisets that contain finitely many distinct formulas. So we do not have a name for the multiset containing one occurrence of every formula in $\mathcal{L}$, but we do have a name for the multiset containing $\aleph_0$ occurrences of $0 = 0$. In a countable language, there can only be countably many multisets of finite base, and cardinality problems are avoided.

I'll use $[A]_n$ to refer to the multiset containing exactly n occurrences of A (and no other formula). I'll let $[A]_\omega$ be the multiset containing $\aleph_0$ occurrences of A, and no other formula.

I'll use $\langle \Gamma \rangle$ for the name for Γ, where Γ is a countable multiset containing at most finitely many distinct formulas.

I'll assume that we can define a function $f(x, y)$ such that $f(n, \langle A \rangle) = \langle [A]_n \rangle$ for all $n < \omega$, and that $f(\omega, \langle A \rangle) = [A]_\omega$.

Let our validity predicate $Val(x, y)$ take names for multisets as the first argument, and names for formulas as the second argument.

Let $Val(x, y)$ obey the following rules, where Γ and Δ are countable multisets containing at most finitely many distinct formulas:

Validity Proof:

$$\frac{\Gamma \vdash \Delta}{\vdash Val(\langle \Gamma \rangle, \langle \Delta \rangle)} \text{ (VP)}$$

Validity Detachment:

$$\frac{}{\Gamma, Val(\langle \Gamma \rangle, \langle \Delta \rangle) \vdash \Delta} \text{ (VD)}$$

Given these rules, we run the risk of paradox. Normally, in a noncontractive logic, we understand $\Gamma, \Gamma' \vdash \Delta$ as a way of writing $\Gamma \boxplus \Gamma' \vdash \Delta$, where $\Gamma \boxplus \Gamma'$ is the multiset such that A occurs in $\Gamma \boxplus \Gamma'$ the sum of the number of times it occurs in Γ and the number of times it occurs in Γ'.

See e.g. [26]: "We form pairwise and countable combinations of multisets using cardinal summation", taking the comma in the premises to reflect this $\boxplus$-style multiset combination.

The problem for noncontractive approaches is that the sum of $\aleph_0$ and 1 is $\aleph_0$, so $[A]_\omega \boxplus \{A\} = [A]_\omega$.

Let $\pi = Val(f(\omega, \langle \pi \rangle), \langle \bot \rangle)$.

We can derive $\vdash \pi$ in the following way. First, we derive $[\pi]_\omega \vdash \bot$:

$$\cfrac{\cfrac{\cfrac{\cfrac{}{[\pi]_\omega, Val(\langle [\pi]_\omega \rangle, \langle \bot \rangle) \vdash \bot} \text{VD}}{[\pi]_\omega, Val(f(\omega, \langle \pi \rangle), \langle \bot \rangle) \vdash \bot} \text{def. of f(x,y)}}{[\pi]_\omega, \pi \vdash \bot} \text{def. of } \pi}{[\pi]_\omega \vdash \bot} [\pi]_\omega \boxplus \pi = [\pi]_\omega$$

From this, we can derive $\vdash \pi$:

$$\cfrac{\cfrac{\cfrac{[\pi]_\omega \vdash \bot}{\vdash Val(\langle[\pi]_\omega\rangle, \langle\bot\rangle)}\text{ VP}}{\vdash Val(f(\omega, \langle\pi\rangle), \langle\bot\rangle)}\text{ def. of } f(x,y)}{\vdash \pi}\text{ def. of } \pi$$

So then we have $\vdash \pi$ and $[\pi]_\omega \vdash \bot$. If we could use infinitely many occurrences of Cut, we'd get $\vdash \bot$. [26] allows for metainferences with countably infinitely many premises (that's how his quantifier introduction rules are defined). So let's assume that we accept the following rule for all multisets Γ:

$$\frac{\Gamma \vdash A \qquad \vdash \gamma_1 \qquad ... \qquad \vdash \gamma_i \qquad ...}{\vdash A}$$

for all $\gamma_i \in \Gamma$ (note that γ_i and γ_{i+1} may be occurrences of the same formula).

For finite Γ, this is a derivable rule using finitely many instances of Cut. I don't see any compelling philosophical reason why it should fail in the infinitary case when it works in finite cases.

But if we have that rule for infinite Γ, then $\vdash \pi$ and $[\pi]_\omega \vdash \bot$ gives us $\vdash \bot$.

That is, as long as we can use $\vdash \pi$ as many times as we'd like. But if VD is valid in the system, then we *proved* $\vdash \pi$ earlier. We didn't just assume it, and so we *should* be able to use it as many times as we want.

And in that case, we have $\vdash \bot$.

3.5 An External Validity Paradox

The above validity curry paradox uses what Mares and Paoli [10] call *internal* consequence. A is an internal consequence of B if $A \vdash B$ is valid. But Mares and Paoli also consider *external* consequence. A is an external consequence of B iff

$$\frac{\vdash A}{\vdash B}$$

is valid.

Although Zardini takes his noncontractive system to allow for infinitely many premises both for internal *and* external consequence, Mares and Paoli explicitly take internal consequence to be finitary. As such, they will not accept the above paradoxes as applying to their chosen logic.

However, Mares and Paoli explicitly take *external* consequence to be infinitary: we can have infinitely many premises when dealing with external consequence.

Mares and Paoli argue that their system should not contain a validity predicate for external consequence. Mares and Paoli appear to reject an external validity predicate in part because it would allow the validity curry paradox to reappear. Mares and Paoli take external validity to contract, and take the validity curry to be solved by giving up contraction. So if we allow for an external validity predicate, there will be no way to block the paradox.

This is not particularly satisfying. They offer no principled reason as to why an external validity predicate is undesirable, save for the fact that it would lead to paradox. But if external validity is a coherent concept that we are using in our logic, then we should be able to model it non-trivially with a formal predicate. Mares and Paoli allow for an *internal* consequence validity predicate, so it seems arbitrary to allow one validity predicate but not the other, given that both notions of consequence are vital to their system.

But the noncontractivist need not deny the coherence of an external validity predicate in order to avoid an externalized validity curry. [24] shows that an external validity curry paradox can only arise in the presence of contraction for metainferences. Namely, the paradox requires that if

$$\frac{\vdash A \qquad \vdash A}{\vdash B}$$

is a valid metainference, then

$$\frac{\vdash A}{\vdash B}$$

is too. But the noncontractivist does not need to claim that this is the case. The noncontractivist only needs to claim that we can contract on the sentences that we have *proven*. [25] argues that if $\vdash A$ is just a hypothetical assumption, rather than a sequent that we have *derived* in the sequent calculus, then the noncontractivist is free to argue that we cannot contract on $\vdash A$ without having to thereby deny that we can use theorems as many times as we'd like. The noncontractivist can maintain that there is a significant difference between

$$\frac{\vdash A \qquad \vdash A}{\vdash B}$$

and

$$\frac{\vdash A}{\vdash B}$$

while still allowing us to contract on theorems. This, combined with the results of [24], means that noncontractivists are free to allow for an external validity predicate *and* allow us to contract on theorems without running the risk of triviality.

Let $EVAL(x, y)$ be a predicate expressing validity for external consequence, which takes as arguments names for multisets with finite root sets. I will write $\langle\Gamma\rangle$ as the name for the multiset Γ. When no confusion can result, I will write $\langle A\rangle$ as the name for the singleton multiset $\langle[A]\rangle$. The predicate should obey the following external variants of (VP) and (VD), where $\Gamma = [\gamma_1, \gamma_2, ...]$:

External Validity Proof: if

$$\frac{\vdash \gamma_1 \qquad \vdash \gamma_2 \qquad ...}{\vdash B}$$

then

$$\frac{}{\vdash EVAL(\langle\Gamma\rangle, \langle B\rangle)}$$

External Validity Detachment:

$$\frac{\vdash EVAL(\langle\Gamma\rangle, \langle B\rangle) \qquad \vdash \gamma_1 \qquad \vdash \gamma_2 \qquad ...}{\vdash B}$$

These rules, taken together with the assumption that external validity always contracts, leads to paradox. Let ϵ be a sentence such that $\epsilon = EVAL(\langle\epsilon\rangle, \langle\bot\rangle)$.

By external validity detachment, we have:

$$\frac{\vdash EVAL(\langle\epsilon\rangle, \langle\bot\rangle) \qquad \vdash \epsilon}{\vdash \bot}$$

But $\epsilon = EVAL(\langle\epsilon\rangle, \langle\bot\rangle)$, so this is equivalent to

$$\frac{\vdash \epsilon \qquad \vdash \epsilon}{\vdash \bot}$$

By external contraction, we have:

$$\frac{\vdash \epsilon}{\vdash \bot}$$

And by external validity proof, we have:

$$\frac{}{\vdash EVAL(\langle\epsilon\rangle, \langle\bot\rangle)}$$

But again, $\epsilon = EVAL(\langle\epsilon\rangle, \langle\bot\rangle)$, so this is equivalent to

$$\frac{}{\vdash \epsilon}$$

We proved earlier that

$$\frac{\vdash \epsilon}{\vdash \bot}$$

And therefore, by external Cut, we have:

$$\frac{}{\vdash \bot}$$

And thus, we have paradox. But notice that $\vdash \epsilon$ had not been *derived* when we contracted on it. It was, in effect, just an external assumption in the valid metainference that is external validity detachment. If we reject contraction for hypothetical assumptions that have not been proven, as [24] and [25] point out, then this use of contraction is illegitimate.

However, this does not completely avoid the problem. An external validity paradox arises even if we give up external contraction for unproven assumptions. The paradox does not require contraction; it only requires that external validity allows for infinitely many premises, which Mares and Paoli explicitly take it to do.

Let $[A]_\omega$ be the multiset containing countably infinitely many occurrences of A, and no other formulas. Let η be a sentence such that $\eta = EVAL(\langle[\eta]_\omega\rangle, \langle\bot\rangle)$.[8] We can then derive $\vdash \bot$ even without contraction.

By external validity detachment, we have:

$$\frac{\vdash EVAL(\langle[\eta]_\omega\rangle, \langle\bot\rangle) \qquad \vdash \eta \qquad \vdash \eta \qquad \ldots}{\vdash \bot}$$

But $\eta = EVAL(\langle[\eta]_\omega\rangle, \langle\bot\rangle)$, so this is equivalent to

$$\frac{\vdash \eta \qquad \vdash \eta \qquad \vdash \eta \qquad \ldots}{\vdash \bot}$$

And by external validity proof, we have:

$$\frac{}{\vdash EVAL(\langle[\eta]_\omega\rangle, \langle\bot\rangle)}$$

But again, $\eta = EVAL(\langle[\eta]_\omega\rangle, \langle\bot\rangle)$, so this is equivalent to

$$\frac{}{\vdash \eta}$$

[8]If our naming convention is sufficiently Gödel-like, then we can do with using a recursive function $f(x, y)$ such that $f(n, \langle A\rangle) = \langle[A]_n\rangle$ for all $n < \omega + 1$. For simplicity, I'll just assume that we can define such a sentence directly.

We have therefore proven $\vdash \eta$. This is not a hypothetical assumption, but a derived sequent. We can therefore use it as many times as we want in a proof. And we proved earlier that

$$\frac{\vdash \eta \qquad \vdash \eta \qquad \vdash \eta \qquad \ldots}{\vdash \bot}$$

So therefore, we have:

$$\frac{}{\vdash \bot}$$

This proof does not use contraction on any assumptions. If we allow for infinitary sequent proofs, as Zardini [26] does, then this does not use contraction at all.[9] We can simply append the proof of $\vdash \eta$ infinitely many times, and have a direct proof of $\vdash \bot$ without any use of external contraction. But if we do not allow for infinitary proofs, then the derivation of $\vdash \bot$ only contracts on a derived sequent, $\vdash \eta$. Even noncontractivists allow contraction on theorems that we have proved. That is an important part of the defense of noncontractive solutions mounted by [25], [10], [6], and others. If we cannot use *theorems* as many times as we would like, then it would seem that all hope of recovering mathematics is lost. Mathematicians need to be able to use theorems arbitrarily many times.

3.6 Discussion

None of the above infinitary paradoxes require Contraction in the usual sense. Contraction is in some sense unavoidable, due to simple facts about transfinite arithmetic: $\aleph_0 + 1 = \aleph_0$. This is most obviously a problem for Zardini's system, presented in [26], because Zardini's system explicitly allows for countably infinite multisets of premises, and his multiplicative quantifiers require metainferences with infinite premises. This is also a problem for the

[9]Note that Zardini's multiplicative quantifiers *require* infinitely many premises in their introduction rule, so this feature cannot easily be removed from his system.

system defended by Mares and Paoli in [10], since their system allows for infinitely many premises when dealing with external consequence.

The paradoxes above are a general problem facing all noncontractive approaches. When dealing with infinitely many premises—as we do when we ask about the consequences of ZFC, and the like—giving up the structural rule of Contraction cannot avoid paradox. Any noncontractive approach that aims to provide a uniform solution to semantic paradoxes must be able to avoid these paradoxes, and explain why they do not pose a problem for the approach.

One option is to argue that an infinitary logic ought to use *sequences*, rather than multisets. We could thereby avoid infinitary contraction by allowing sequences of any countable ordinal. For although the cardinal $\aleph_0 = \aleph_0 + 1$, this is not true for ordinals: $\omega \neq \omega + 1$. However, this will not prevent infinitary paradoxes. For even in ordinal arithmetic, $1 + \omega = \omega$. As such, as long as we have the structural rule of Permutation:

$$\frac{\Gamma, A, B, \Gamma', \vdash \Delta}{\Gamma, B, A, \Gamma' \vdash \Delta}\text{Permutation}$$

the paradoxes will reappear. We can use the infinitary liar paradox as an example. Suppose that we take premises and conclusions to be *sequences*, rather than multisets. Given a formula A and a countable ordinal α, let $\{A\}_\alpha$ be the α-length sequence of occurrences of A. Let $\oplus$ be a concatenation operator, such that $\Gamma \oplus \Delta$ is the sequence consisting of all of the γs in Γ followed by all of the δs in Δ. When no confusion can arise, I will use $A \oplus \Gamma$ for $\{A\}_1 \oplus \Gamma$. Note that $A \oplus \{A\}_\omega = \{A\}_\omega$.

Suppose again that we have a conjunction $\bigwedge$, which obeys the following rules:

$$\frac{\Gamma \vdash \Delta}{\bigwedge \Gamma \vdash \Delta}\bigwedge\text{L}$$

$$\frac{\Gamma_1 \vdash A_1, \Delta_1 \qquad \Gamma_2 \vdash A_2, \Delta_2 \qquad ... \qquad \Gamma_n \vdash A_n, \Delta_n \qquad ...}{\Gamma_1, \Gamma_2, ...\Gamma_n ... \vdash \bigwedge[A_i], \Delta_1, \Delta_2, ...\Delta_n ...}\bigwedge\text{R}$$

Where here Γs and Δs are sequences. Note in the presence of Cut, $\bigwedge$ will obey the following introduction rule:

$$\frac{A, \bigwedge \Gamma \vdash \Delta}{\bigwedge (A \oplus \Gamma) \vdash \Delta} \bigwedge\text{-intro}$$

For all $\gamma \in \Gamma$, we have:

$$\gamma \vdash \gamma$$

Thus:

$$\frac{\gamma \vdash \gamma \qquad ...}{\Gamma \vdash \bigwedge \Gamma}$$

Thus by Cut, we have:

$$\dfrac{\dfrac{A, \bigwedge \Gamma \vdash \Delta \qquad \Gamma \vdash \bigwedge \Gamma}{A, \Gamma \vdash \Delta}}{\bigwedge (A \oplus \Gamma) \vdash \Delta}$$

Given such a conjunction, we will get triviality without Contraction even if we take premises and conclusions to be sequences rather than multisets. Let L the sentence such that $\langle L \rangle$ is the name for the sentence $\bigwedge \{\neg T \langle L \rangle\}_\omega$. The derivation is analogous to the multiset version given in section 3.3; only Part A is different:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{L \vdash L}{L \vdash T\langle L \rangle}\text{T-intro}}{L, \neg T\langle L \rangle \vdash}\neg\text{L}}{\neg T\langle L \rangle, L \vdash}\text{Permutation}}{\neg T\langle L \rangle, \bigwedge \{\neg T\langle L \rangle\}_\omega \vdash}\text{def. of } L}{\bigwedge (\neg T\langle L \rangle \oplus \{\neg T\langle L \rangle\}_\omega) \vdash}\bigwedge\text{-intro}}{\bigwedge \{\neg T\langle L \rangle\}_\omega \vdash}\neg T\langle L \rangle \oplus \{\neg T\langle L \rangle\}_\omega = \{\neg T\langle L \rangle\}_\omega}{L \vdash}\text{def. of } L$$

Parts B and C are exactly the same, with only minor notation variations. The other infinitary paradoxes can be similarly adapted to a sequence setting.

Note that in the above derivation, Permutation is applied only to sentences; we are not permuting transfinite sequences, which may be considered problematic, as it involves infinitely many syntactic steps. The assumption that we can permute adjacent sentences in a sequence seems to be a very minor assumption, and one that is almost certainly safe in many contexts. It seems very implausible that the order in which you list premises can change what follows from those premises. For the noncontractive approach to avoid infinitary paradoxes, it is therefore not enough to argue that an infinitary logic must involve infinite *sequences*, rather than infinite *multisets*.

Another option is simply to deny the coherence of infinitary conjunction or an infinitary Val-predicate rules. But such a restriction on vocabulary seems hopelessly ad hoc, and threatens to undermine the noncontractivist's claim of a uniform solution. Part of the appeal of substructural logics is supposed to be that, unlike more traditional nonclassical approaches that have to adjust the rules for different connectives to solve different paradoxes, substructural logics can solve the paradoxes in one stroke. The fact that paradox can arise with infinite conjunction and no Val predicate, and also arise with a Val predicate and no infinite conjunction, suggests that neither piece of vocabulary is really the underlying problem. So solving the paradox by restricting either or both pieces of vocabulary would fail to provide a uniform diagnosis of the infinitary paradoxes, and instead make only ad hoc vocabulary restrictions. And besides, neither piece of vocabulary seems particularly problematic. Infinite conjunctions are perfectly coherent, and are often used in mathematics, such as in the usual presentation of Vaught's conjecture.[10] And once you've allowed for valid infinitary inferences, your validity predicate should allow you to talk about them. There seems to be no principled reason to reject infinitary conjunctions or infinitary validity

[10]See e.g.[2], [1], [23], and [11].

predicates once we have allowed arguments with infinitely many premises.

The more promising option for the noncontractivist is to forbid arguments with infinitely many premises, and to insist that we only ever consider inferences with finitely many premises. The noncontractivist can claim that they are only interested in inferences as proof-theoretic objects, meaning that an inference $\Gamma \vdash \Delta$ indicates that there exists a proof taking the γs as assumptions that has (one of) the δs as its conclusion. Since a proof can only contain finitely many steps, inferences with infinitely many premises are simply beside the point. As a result, the noncontractivist may argue that these apparent paradoxes aren't really paradoxes for their noncontractive systems, because these paradoxes involve a completely unrelated notion of inference.

In many contexts, when we are interested in the *exact* premises that we need in order to prove a conclusion, we may not need or want infinite multisets of premises. In linear logic, which does not include Weakening as a rule, this may be the question we are interested in.[11] No proof ever has infinitely many premises, so no infinite set of premises is the *exact* set of premises needed for a proof.

However, there are certainly some contexts in which we want to ask whether or not the premises we have are *sufficient* to prove something. This is the question that Cintula and Paoli take affine logic (linear logic + weakening) to be suitable for ([6]; fn. 7).

If we allow for weakening, and we are interested in what our premises are *sufficient* to prove, then it seems arbitrary to restrict to finite multisets of premises. We are often interested in knowing what an infinite set of premises can be used to prove. ZFC, for example, is not finitely axiomatizable. But we certainly want to ask what those axioms are sufficient to prove. And when we ask that question, we are not assuming that we can have infinitely long proofs; we are instead interested in what finite proofs can be constructed using the resources in that infinite set of axioms.

[11] For more on linear logic, see [9] and [22].

Defenders of the noncontractive approach may argue that they can give an indirect account of the consequences of an infinite collection of axioms as follows: A is a consequence of the infinite collection Γ iff $\Delta \vdash A$ for some finite $\Delta \subset \Gamma$. Let's use $\rhd$ as a symbol for this notion: $\Gamma \rhd A$ iff there exists a finite sub-multiset Δ of Γ such that $\Delta \vdash A$.

Unfortunately, the paradox will reappear as soon as we move to reasoning with $\rhd$. Let $INFVAL(x, y)$ be a predicate representing $\rhd$. If this predicate is going to represent $\rhd$, then by analogy to Val and $\vdash$, $INFVAL$ should obey the following rules:

Infinitary Validity Proof:

$$\frac{\Gamma \rhd \Delta}{\rhd Val(\langle\Gamma\rangle, \langle\Delta\rangle)} \text{ (IVP)}$$

Infinitary Validity Detachment:

$$\frac{}{\Gamma, Val(\langle\Gamma\rangle, \langle\Delta\rangle) \rhd \Delta} \text{ (IVD)}$$

Let $\pi = INFVAL(f(\omega, \langle\pi\rangle), \langle\bot\rangle)$.

We can derive $\rhd \pi$ in the following way. First, we derive $[\pi]_\omega \rhd \bot$:

$$\cfrac{\cfrac{\cfrac{\cfrac{}{[\pi]_\omega, Val(\langle[\pi]_\omega\rangle, \langle\bot\rangle) \rhd \bot} \text{ IVD}}{[\pi]_\omega, Val(f(\omega, \langle\pi\rangle), \langle\bot\rangle) \rhd \bot} \text{ def. of f(x,y)}}{[\pi]_\omega, \pi \rhd \bot} \text{ def. of } \pi}{[\pi]_\omega \rhd \bot} \; [\pi]_\omega \boxplus \pi = [\pi]_\omega$$

From this, we can derive $\rhd \pi$:

$$\cfrac{\cfrac{\cfrac{[\pi]_\omega \rhd \bot}{\rhd Val(\langle[\pi]_\omega\rangle, \langle\bot\rangle)} \text{ IVP}}{\rhd Val(f(\omega, \langle\pi\rangle), \langle\bot\rangle)} \text{ def. of } f(x,y)}{\rhd \pi} \text{ def. of } \pi$$

Given the definition of $\rhd$, $\rhd \pi$ entails $\vdash \pi$. And we also have $[\pi]_\omega \rhd \bot$, which means that there is a finite sub-multiset Γ of $[\pi]_\omega$ such that $\Gamma \vdash \bot$. Γ must be a multiset containing finitely many instances of π, so from $\Gamma \vdash \pi$ and $\vdash \pi$ we need only finitely many applications of Cut to get $\vdash \bot$.

Therefore even if the noncontractivist aims to give an indirect characterization of what it means to be a consequence of an infinite collection of premises, paradox will once again rear its ugly head. This means that, for the noncontractivist to give a uniform solution to semantic paradoxes, they must deny that there is *any* usable notion of consequence according to which we can have infinitely many premises, or can contract on premises.

The problem then is that, as a simple matter of empirical fact, we are often interested in the consequences of infinite collections of premises. As long as we are interested in the consequences of infinite collections of axioms (and mathematics does not seem to be losing interest in that question any time soon) we should be able to formulate a consequence relation that tolerates infinitely many premises. And once we do that, noncontractive approaches are no longer sufficient to avoid paradoxes.

Thus to avoid the paradoxes here, the noncontractivist would have to say that there is *no coherent sense* to be made of infinitely many premises. But given that many (perhaps most) of the axiom sets that interest mathematicians are infinite, this is untenable. This means that there are notions of consequence that the noncontractive approach to paradox simply ignores. As such, the noncontractive approach does not provide a uniform solution to semantic paradoxes. It may solve the paradoxes *for specific contexts*—contexts in which we are interested in resource-conscious proofs—but it does not solve the paradoxes *in general.* The paradoxes still arise in cases where we are interested in the consequences of infinite collections of axioms, and the noncontractive approach does nothing to solve the paradoxes in those contexts.

3.7 Contraction and Closure

The above paradoxes show that noncontractive approaches cannot allow us to ask about the consequences of infinite collections of non-logical axioms. This is a symptom of a general problem for noncontractive approaches: a noncontractive logic cannot tell us what the consequences of a collection of axioms are, because we must make a choice about which proofs to use the axioms in. We can only use an instance of an axiom once, and we must choose how we use it. Given an instance of $A \wedge B$, we must choose whether to infer A or to infer B from it; we cannot infer both A and B from a single occurrence of $A \wedge B$, regardless of whether $\wedge$ is additive or multiplicative. As such, there is no fact about what *the* consequences of our axioms are; they can only give us some options from which to choose.

This problem arises because noncontractive consequence relations are not closure operators. A closure operator C(X) is an operator on sets (or multisets) such that:

1. $X \subseteq C(X)$
2. if $X \subseteq Y$ then $C(X) \subseteq C(Y)$
3. $C(X) = C(C(X)$

Ordinarily, a (single-conclusion) consequence relation $\vdash$ corresponds to a closure operator $C(X)$, such that $\Gamma \vdash A$ iff $A \in C(\Gamma)$. This give us the usual properties that we want:

(1) gives us identity: $A \vdash A$

(2) gives us monotonicity: $\Gamma \vdash A$ implies $\Gamma, B \vdash A$

(3) gives us "Cautious Cut": $\Gamma \vdash A$ and $\Gamma, A \vdash B$ implies $\Gamma \vdash B$.

But giving up Cut or Contraction for a consequence relation means that we cannot have closure. (3) gives us Cut, automatically. And (1), (2) and (3) together give us Contraction:

Suppose $\Gamma, A, A \vdash B$.

By (1), $A \vdash A$.

By (2), $\Gamma, A \vdash A$.

By (3), $\Gamma, A \vdash B$.

In [16], Ripley shows that any consequence relation that corresponds to a closure operator and validates Cut must also validate Contraction. That means that noncontractive consequence relations cannot correspond to closure operators.

This is a serious problem, particularly if we want to use our logic to add non-logical axioms and find out what they entail. When we want to discover the consequences of a collection of non-logical axioms like ZFC, what we want is a closure operator. Normally can take e.g. the axioms of ZFC, and know that there is a singular set of theorems provable from those axioms. This is the *closure* of the axioms.

This is what allows us to use a collection of axioms to present an entire theory. As Beall puts it in [3]. "Give to logic your theory Γ, and then sit back: logic 'freely' or 'automatically' expands your theory to $C(\Gamma)$, which contains all of Γ's (singleton) consequences."

This power of a logic is dependent on it corresponding to a closure operator. Without a closure operator, we cannot use a logic to present the entirety of a theory using just a set of axioms, and we cannot use a logic to discover the consequences of a collection of non-logical axioms.

Classical logic, paracomplete logics, and paraconsistent logics all do this: they each correspond to a closure operator, and therefore we can use these logics to discover the consequences of a collection of axioms or assumptions in the usual way. But noncontractive logics do not allow us to do this.

However, we are often concerned with discovering what *the* consequences of a collection of assumptions or axioms are. Semantic paradoxes arise in those contexts. Noncontractive logics cannot be used in those contexts, and so the noncontractive approach does nothing to solve the paradoxes in those contexts.

3.8 Conclusion

Noncontractive solutions to semantic paradoxes aim to provide a single uniform solution that solves every semantic paradox by blocking Contraction. Even if this approach works well for inferences with finitely many premises, it fails to solve paradoxes involving inferences with infinitely many premises. In order to avoid such paradoxes, the noncontractivist would have to deny that there is any coherent sense to be made of an inference with infinitely many premises. But denying the coherence of inferences with infinitely many premises is problematic, because not every theory we want to talk about is finitely axiomatizable. And even if every theory were finitely axiomatizable, noncontractive logics still cannot give us a single theory that is entailed by a collection of axioms: it can only give us some mutually exclusive options of what we *could* prove from those axioms.

Much of mathematical research involves discovering the consequences of infinite collections of axioms like ZFC. The noncontractive approach does not allow us to discuss the consequences of infinite collections like the ZFC axioms. If we can't talk about what follows from an infinite (multi)set of premises, then it's not clear how we can do anything approaching contemporary mathematics. So even if noncontractivists are willing to revise their systems of logic to block any and all occurrences of infinite collections of premises, this would fail to provide a uniform solution to paradoxes. We are often interested in the consequences of infinite collections of axioms. If noncontractive logics cannot be applied in those contexts, then noncontractive logics simply do not solve the paradoxes as they arise in those contexts. Rather than providing a uniform solution to paradoxes wherever they arise, the noncontractive approach simply ignores many contexts in which we need a solution to semantic paradoxes.

Bibliography

[1] Ackerman, Nathanael Leedom (2015). Vaught's Conjecture Without Equality. *Notre Dame Journal of Formal Logic*, 56 (4), 573-582.

[2] Baldwin, John T., Friedman, Sy D., Koerwien, Martin, and Laskowski, Michael C. (2016). Three Red Herrings Around Vaught's Conjecture. Transactions of the American Mathematical Society, 368, 3673-3694.

[3] Beall, JC (2015) "Free of Detachment: Logic, Rationality, and Gluts" *Nous* 49(2): 410-23.

[4] Beall, JC and Murzi, Julien (2013). Two Flavors of Curry's Paradox. *Journal of Philosophy*, 110 (3), 143-165.

[5] Caret, Colin R. and Weber, Zach (2015). A Note on Contraction-Free Logic for Validity. *Topoi*, 34 (1): 63-74.

[6] Cintula, Petr and Paoli, Francesco (2016). Is Multiset Consequence Trivial? *Synthese*. https://doi-org.ezproxy.gc.cuny.edu/10.1007/s11229-016-1209-7.

[7] Da Ré, Bruno and Rosenblatt, Lucas (2018). Contraction, Infinitary Quantifiers, and Omega Paradoxes. *Journal of Philosophical Logic*, 47 (4), 611-629.

[8] French, Rohan and Ripley, David (2015). Contractions of Noncontractive Consequence Relations. *The Review of Symbolic Logic*. 8 (3), 506-528.

[9] Girard, Jean-Yves (1987). Linear Logic. *Theoretical Computer Science*, 50, 1-102.

[10] Mares, Edwin and Paoli, Francesco (2014). Logical Consequence and the Paradoxes. *Journal of Philosophical Logic*, 43 (2-3), 439-469.

[11] Morley, Michael (1970). The Number of Countable Models. *The Journal of Symbolic Logic*, 35 (1), 14-18.

[12] Paoli, Francesco (2005). The Ambiguity Of Quantifiers. *Philosophical Studies*, 124 (3), 313-330.

[13] Priest, Graham (manuscript). Substructural Solutions to the Semantic Paradoxes: Dialetheism in Sheep's Clothing?

[14] Restall, Greg (1994). *On Logics Without Contraction*. Doctoral dissertation, University of Queensland. Retrieved from https://consequently.org/papers/onlogics.pdf.

[15] Ripley, David (2015a). Comparing Substructural Theories of Truth. *Ergo*, 2, 299-328.

[16] Ripley, David (2015b). Contraction and Closure. *Thought*, 4 (2), 131-138.

[17] Rosenblatt, Lucas (2019a). Non-Contractive Classical Logic. *Notre Dame Journal of Formal Logic*, 60 (4), 559-585.

[18] Rosenblatt, Lucas (2019b). On Structural Contraction and Why It Fails. *Synthese*. https://doi.org/10.1007/s11229-019-02241-6.

[19] Shapiro, Lionel (2011). Deflating Logical Consequence. *Philosophical Quarterly*, 61 (243), 320-342.

[20] Shapiro, Lionel (2013). Validity Curry Strengthened. *Thought*, 2, 100-107.

[21] Shapiro, Lionel (2015). Naive Structure, Contraction and Paradox. *Topoi*, 34 (1), 75-87.

[22] Troelstra, Anne (1992). *Lectures on Linear Logic*, Stanford: CSLI Publications.

[23] Vaught, Robert (1973). Descriptive Set Theory in $L_{\omega_1\omega}$. *Cambridge Summer School in Mathematical Logic*, edited by A.R.D. Mathias and H. Rogers, Lecture Notes in Mathematics, vol. 337, Springer-Verlag, Berling, Heidelberg, and New York, 574-598.

[24] Wansing, Heinrich and Priest, Graham (2015). External Curries. *Journal of Philosophical Logic*, 44: 453-471.

[25] Weber, Zach (2016). On Closure and Truth in Substructural Theories of Truth. *Synthese*. https://doi-org.ezproxy.gc.cuny.edu/10.1007/s11229-016-1226-6.

[26] Zardini, Elia (2011). Truth Without Contra(di)ction. *Review of Symbolic Logic*, 4 (4), 498-535.

[27] Zardini, Elia (2013a). Naive Logic Properties and Structural Properties. *Journal of Philosophy*, 110 (11), 633-644.

[28] Zardini, Elia (2013b). Naive Modus Ponens. *Journal of Philosophical Logic*, 42 (4): 575-593.

[29] Zardini, Elia (2016). Restriction by Noncontraction. *Notre Dame Journal of Formal Logic*, 57(2):287-327.

[30] Zardini, Elia (2019). Instability and Contraction. *Journal of Philosophical Logic*, 48:155-188.

Chapter 4

An Infinite Hierarchy of Validity Curry Paradoxes

4.1 Introduction

A substructural logic is any logic that gives up one of the usual *structural* rules, usually either Cut or Contraction:[1]

$$\frac{\Gamma, A \vdash \Delta \qquad \Gamma' \vdash A, \Delta'}{\Gamma, \Gamma' \vdash \Delta, \Delta'}\,\text{Cut}$$

$$\frac{\Gamma, A, A \vdash \Delta}{\Gamma, A \vdash \Delta}\,\text{Contraction}$$

Logics giving up Cut or Contraction are called *nontransitive* and *noncontractive*, respectively. Substructural approaches aim to provide a uniform solution to semantic paradoxes.[2] Each semantic paradox involves either Cut or Contraction, and so supporters of substructural approaches argue that rejecting one of these rules allows us to block all the usual paradoxical arguments at the exact same step.

If successful, these substructural approaches can give us a single, uniform solution that applies widely to all of the semantic paradoxes. By comparison, more traditional nonclassical approaches appear somewhat *ad hoc*: paracomplete and paraconsistent approaches must give up one rule to block the Liar Paradox, another to block Curry's Paradox, and yet another to block what Beall and Murzi [4] call the *validity curry* paradox.

The validity curry paradox is particularly problematic for fully structural approaches, because the paradox uses only the structural rules of Cut and Contraction, plus very plausible rules governing a validity predicate. This paradox has been presented as a consideration in favor of substructural approaches to semantic paradoxes.[3] While paracomplete and paraconsistent approaches can tolerate a naive theory of truth, it seems that only substructural logics can provide a naive theory of validity.

[1]Although see [20] for a substructural logic that gives up Identity instead.

[2]For recent defenses of noncontractive approaches, see [42], [41], [27], [28], [23], [5], [38], [37], and [36]. For recent defenses of nontransitive approaches, see [8], [6], [34], [33], [32], and [24].

[3]See, for example, [33], [32], [42], [4], and [36].

The validity curry paradox involves a validity predicate that is meant to represent the validity relation between (multisets of) formulas. But those are not the only sorts of inferences that we're interested in. We are also interested in *metainferences*: inferences between inferences.

For example, Cut and Contraction are both metainferences: they take an inference as a premise, and have an inference as a conclusion. To even be able to discuss whether or not to move to a substructural logic, we need to be able to discuss which metainferences are valid. But even in a classical setting, we need to be able to talk about metainferences. For example the deduction theorem, which is often presented as something like "if $A \vdash B$, then $\vdash A \to B$" can be understood as a metainference:

$$\frac{A \vdash B}{\vdash A \to B}$$

Metainferences are also useful in comparing features of different logics. For example, intuitionistic logic has the property that if $\vdash AVB$, then either $\vdash A$ or $\vdash B$. Classical logic does not have this property because, for example, $\vdash p \vee \neg p$ is valid but neither $\vdash p$ nor $\vdash \neg p$ is.

One way to understand this difference is in terms of metainferences: the multi-conclusion metainference

$$\frac{\vdash A \vee B}{\vdash A \qquad \vdash B}$$

is valid in intuitionistic logic, but not in classical logic.

As long as we are in the business of representing inferences that we are interested in discussing, as the validity curry paradox assumes, it stands to reason that we should be able to represent *meta*inferences, not just first-order inferences. However, a higher-order version of the validity curry paradox reappears at the metainferential level. Providing a uniform solution to the validity curry paradoxes of different orders proves to be much more challenging than solving the original validity curry paradox alone.

There has already been some work done on the difficulties of representing metainferential validity. For example, in [29] introduces the concept of *internalization*, which is one way for a logic to represent the validity-preservation of its own metarules. Rosenblatt argues that substructural theories cannot internalize their own metarules. In [3], Barrio, Rosenblatt and Tajer show that STV, the nontranstivie strict-tolerant logic with the validity rules VP and VD, cannot internalize all of its own metarules, and that the most obvious way to strengthen the validity rules allows the theory to (wrongly) internalize metarules that *don't* preserve validity. In [21], and [22], Hlobil considers an admissibility version of the validity curry paradox, and argues that the nontrasitive approach should reject the rule VD in order to better internalize its own metarules. Smilarly, the infinite-premise "external validity curry" that I introduced in "Infinite-Premise Paradoxes and Contraction" is a sort of metainferential curry paradox.

In this paper, I show that this difficulty in representing valid metainferences generalizes beyond the internalization of metarules, or the problems of infinite premises. Even if we add a separate validity predicate to represent metainferential validity, and allow only finitely many premises, the validity curry paradox will reappear at the metainferential level.

In this paper I present metainferential versions of the validity curry paradox, and show that the validity curry phenomenon appears at every level of metainference. The metainferential validity curry does not involve Cut or Contraction; it instead involves higher level analogues that are validated by current substructural logics like those of [6] and [42]. But the problems do not end there: the problem reappears at the level of metametainferences, and metametametainferences, and so on ad infinitum. I argue that this poses a problem for substructural approaches to semantic paradoxes, and that in particular it threatens to undermine their ability to provide a uniform solution to the semantic paradoxes. It also poses a problem for non-substructural approaches, and may force us to rethink what it means to provide a uniform solution to these paradoxes.

In section 4.2, I present the notation and metainferential framework that I will use throughout the paper. In section 4.3, I discuss the validity curry paradox and the substructural solutions to it. In section 4.4, I introduce a metainferential validity curry paradox, and show that it does not involve the usual Cut or Contraction rules. In section 4.5, I introduce an infinite hierarchy of metainferential validity curry paradoxes, none of which involve the Cut or Contraction rules of any lower level. In section 4.6, I discuss what it means to provide a uniform solution to the validity curry paradoxes in the hierarchy. In sections 4.9 and 4.10, I argue that giving up either of VP or VD at every level is untenable. In sections 4.7 and 4.8 I argue that giving up Cut or Contraction at every level of inference is worse than giving up either rule at only the first level. In section 4.11, I reconsider what it means to provide a uniform solution, and argue that the non-classical logician can reject VP in *some* cases and VD in others while still providing a uniform solution. In section 4.12, I conclude that the metainferential hierarchy of validity curry paradoxes poses new problems all approaches to semantic paradox, but that we do not need to move to a substructural logic to provide a uniform solution. I close with some remarks about the future of metainferential paradoxes.

4.2 Notation

When introducing the validity curry paradox in [4], Beall and Murzi consider a validity predicate that takes as arguments names for formulas, such as $\langle A \rangle$ for the formula A. However, we are often interested in arguments with multiple premises and multiple conclusions. We could always conjoin premises and disjoin conclusions to get one premise and one conclusion, so that e.g. $A, B \vdash C, D$ becomes $A \wedge B \vdash C \vee D$. But this will not work in all cases. In supervaluationist logics, $\Gamma \vdash A, B$ is not equivalent to $\Gamma \vdash A \vee B$. And in subvaluationist logics,

$A, B \vdash \Delta$ is not equivalent to $A \wedge B \vdash \Delta$. In such cases, the one-premise-and-one-conclusion format will fail to capture all of the inferences that we want to capture.

With that in mind, it would be better to have names for finite multisets of formulas.[4] I will take $\langle\Gamma\rangle$ to be the name for the multiset Γ, take e.g. $\langle[A, A, B, ...]\rangle$ to be the name for the multiset $[A, A, B...]$ and $\langle[A]\rangle$ to be the name for the singleton multiset $[A]$. When no confusion can arise, I will drop the square brackets [ ] and just write e.g. $\langle A, A, B\rangle$.

In addition to names for multisets, we'll need a way to talk about metainferences, or inferences between inferences. To do that, we will need names not only for multisets, but for *inferences*. To do this, we can think of inferences as ordered pairs of multisets of formulas: the inference $\Gamma \vdash \Delta$ is simply the ordered pair (Γ, Δ). We can similarly think of (single-premise and single-conclusion) metainferences as ordered pairs of ordered pairs of multisets, and metametainferences as ordered pairs of ordered pairs of ordered pairs of multisets, and so on. As long as we restrict ourselves to finite multisets, a countable language is free to contain names for every finite multiset of formulas, and every ordered pair of finite multisets, and every ordered pair of ordered pairs of finite multisets, and so on, without increasing the cardinality of the language. We can do this via gödel numbering, or by extending the language with a set of distinguished names for multisets of formulas and ordered pairs of ... ordered pairs of multisets of formulas. In what follows, I will assume that our language contains such names.

In order to discuss metainferences, and metametainferences, and metametametainferences, and so on, we will need a way to represent meta^{n}inferences of arbitrary level n. For this purpose, we can use higher-level sequents, which take sequents of lower levels as premises. This was pioneered by Kosta Došen, who introduced higher-level sequent systems for use in modal logic and other purposes.[5] For my purposes in this paper, I will specifically be using

[4] If we included names for infinite multisets, then the cardinality of our set of terms would be greater than the cardinality of our language, which would be a problem.

[5] See in particular [10], [11], [12], and [13].

tools developed in [1], with some terminology from [24] and [35].

Let $\mathcal{L}$ be a language, which contain names for finite multisets of formulas, ordered pairs of finite multisets of formulas, ordered pairs of ordered pairs of finite multisets of formulas, and so on.

I'll call $\mathcal{L}$-formulas *0-inferences*.

Let a *1-inference* be an ordered pair of finite multisets of $\mathcal{L}$-formulae: $\langle \Gamma, \Delta \rangle$. I will write this as $\Gamma \Rightarrow_1 \Delta$, and I will write the name for $\Gamma \Rightarrow_1 \Delta$ as $\langle \Gamma \Rightarrow_1 \Delta \rangle$.

1-inferences are the usual inferences between sets of formulae, like modus ponens:

$$A, A \rightarrow B \Rightarrow_1 B$$

Throughout this paper, if the 1-inference $\Rightarrow A$ is valid, I will call A a "theorem" of the logic.

But we also want to consider inferences between inferences, or *metainferences*, such as Cut and Contraction:

$$\frac{\Gamma, A \Rightarrow_1 \Delta \qquad \Gamma' \Rightarrow_1 A, \Delta'}{\Gamma, \Gamma' \Rightarrow_1 \Delta, \Delta'} \text{Cut}$$

$$\frac{\Gamma, A, A \Rightarrow_1 \Delta}{\Gamma, A \Rightarrow_1 \Delta} \text{Contraction}$$

To do this, we will look not only at 1-inferences, but at *2-inferences*. Let a 2-inference be an ordered pair of multisets of 1-inferences. I will write this as $\Gamma \Rightarrow_2 \Delta$, where Γ and Δ are sets of 1-inferences rather than sets of formulae.

On this notation, Cut becomes:

$$\{\Gamma, A \Rightarrow_1 \Delta; \Gamma' \Rightarrow_1 A, \Delta'\} \Rightarrow_2 \{\Gamma, \Gamma' \Rightarrow_1 \Delta, \Delta'\}$$

And Contraction becomes:

$$\{\Gamma A, A \Rightarrow_1 \Delta\} \Rightarrow_2 \{\Gamma, A \Rightarrow_1 \Delta\}$$

These are the same metainferences that we've always used, just presented in a different notation. We can use this notation to discuss metainferences of arbitrary levels. For any n, we can use this notion to consider meta^{n}inferences.

Importantly, this gives us a way to construct proofs that a metainference is valid, analogous to the way in which the usual sequent calculus allows us to construct a proof that an inference is valid.

For all $n < \omega$, let an *n+1-inference* be an ordered pair of finite multisets of n-inferences. 1-inferences are inferences between formulas, 2-inferences are the metainferences, 3-inferences are metametainferences, etc.

For any $n > 0$, we can write an n-inference as $\Gamma \Rightarrow_n \Delta$, where Γ and Δ are finite multisets of (n-1)-inferences.

This gives us an infinite hierarchy of meta^{n}inferences. At any level of inference n, we can use this notation to talk about which n-inferences are valid.

4.2.1 A Note on Metainferential Validity

Given a class of models, we can define a notion of validity for 1-inferences by defining a *satisfaction* relation between models and inferences. For example, we might define the (single-conclusion) classical logic consequence relation $\vdash_{CL}$ over the class of boolean valuations $\mathfrak{M}$ as follows:

$$\Gamma \vdash_{CL} A \text{ iff for all } v \in \mathfrak{M}, \text{ either } v(\gamma) = 0 \text{ for some } \gamma \in \Gamma, \text{ or } v(A) = 1$$

A classical counterexample to an inference $\Gamma \vdash A$ is a boolean valuation that assigns 1 to all of the premises, and assigns 0 to the conclusion. A valuation satisfies an inference if it is not a counterexample. A classically valid 1-inference is an inference that has no such counterexamples, or equivalently, an inference that is satisfied at every model.

Given a model-theoretically defined notion of validity for 1-inferences, we can define at least two notions of validity for 2-inferences: *global* and *local*:

Definition 4.2.1. *Local Metainferential Validity*

A metainference

$$\frac{\Gamma \vdash \Delta}{\Sigma \vdash \Pi}$$

is locally valid iff at every model m*, if* $\Gamma \vdash \Delta$ *is satisfied at* m*, then* $\Sigma \vdash \Pi$ *is satisfied at* m*.*

Definition 4.2.2. *Global Metainferential Validity*

A metainference

$$\frac{\Gamma \vdash \Delta}{\Sigma \vdash \Pi}$$

is globally valid iff if (materially) $\Gamma \vdash \Delta$ *is valid (satisfied at all models), then* $\Sigma \vdash \Pi$ *is valid.*

These are both model-theoretic notions of metainferential validity. But there are also two closely related proof-theoretic notions that we might be interested in:

Definition 4.2.3. *Derivable Metainferential Rule*

A metainference

$$\frac{\Gamma \vdash \Delta}{\Sigma \vdash \Pi}$$

is a derivable rule iff there exists a proof from $\Gamma \vdash \Delta$ *to* $\Sigma \vdash \Pi$*.*

Definition 4.2.4. *Admissible Metainferential Rule*

A metainference

$$\frac{\Gamma \vdash \Delta}{\Sigma \vdash \Pi}$$

is an admissible rule iff if (materially) there exists a proof of $\Gamma \vdash \Delta$*, then there exists a proof of* $\Sigma \vdash \Pi$.

In general, given a proof theory and a (model-theoretic) semantics for which it is sound and complete, global validity and admissibility coincide. But local validity and derivable rules do not necessarily coincide.[6]

In [1], Barrio, Pailos and Szmuc construct a metainferential hierarchy of strict-tolerant logics using the *local* notion of metainferential validity, defined over 3-valued Strong Kleene models. Their hierarchy is defined in terms of local metainferential validity, but the metainferential arguments that I introduce in this paper could be run using any of these four notions (and any variations thereof).

In what follows, I will assume that the notion of metainferential validity in play is the *global* notion of metainferential validity. This is the more common notion of metainferential validity; it is the notion involved in the deduction theorem in classical logic, and the Necessitation rule in modal logic. It is also clearly the notion of metainferential validity used in the derivation of the validity curry paradox; as we will see, the rule Validity Proof (VP) is only desirable as a global metainference.

4.3 The Validity Curry Paradox

In [4], Beall and Murzi present a validity curry paradox that involves a validity predicate $Val(x, y)$, governed by the following rules:[7]

Validity Proof:

[6]See [9] for a further discussion of the relationships between these notions of metainferential validity.
[7]Notation changed to match the notation used in this paper.

$$\frac{\Gamma \Rightarrow \Delta}{\Rightarrow Val(\langle \Gamma \rangle, \langle \Delta \rangle)} \text{ (VP)}$$

Validity Detachment:

$$\frac{}{\Gamma, Val(\langle \Gamma \rangle, \langle \Delta \rangle) \Rightarrow \Delta} \text{ (VD)}$$

These rules lead to triviality in the presence of the structural rules of Cut and Contraction:[8]

$$\frac{\Gamma, A \Rightarrow \Delta \qquad \Gamma' \Rightarrow A, \Delta'}{\Gamma, \Gamma' \Rightarrow \Delta, \Delta'} \text{ Cut}$$

$$\frac{\Gamma, A, A \Rightarrow \Delta}{\Gamma, A \Rightarrow \Delta} \text{ Contraction}$$

Let π be a sentence that is equivalent to $Val(\langle \pi \rangle, \langle \bot \rangle)$. Using this sentence, we can use Cut, Contraction, and the validity rules to derive $\Rightarrow \bot$. Thus any logic validating all four rules is trivial. The derivation has two parts: I'll call them Part A, and Part B. In general, noncontractive approaches object to Part A, while nontransitive approaches object to Part B.[9]

Part A is a derivation of $\pi \Rightarrow \bot$, using only (VD) and Contraction:

$$\cfrac{\cfrac{\cfrac{}{\pi, Val(\langle \pi \rangle, \langle \bot \rangle) \;\Rightarrow\; \bot} \text{ (VD)}}{\pi, \pi \;\Rightarrow\; \bot} \text{ (def. of } \pi\text{)}}{\pi \;\Rightarrow\; \bot} \text{ (Contraction)}$$

Part B uses Part A, along with (VP) and Cut, to derive $\Rightarrow \bot$:

$$\cfrac{\cfrac{\cfrac{\cfrac{[\text{Part A}]}{\pi \;\Rightarrow\; \bot}}{\Rightarrow\; Val(\langle \pi \rangle, \langle \bot \rangle)} \text{ (VP)}}{\Rightarrow\; \pi} \text{ (def. of } \pi\text{)} \qquad \cfrac{[\text{Part A}]}{\pi \;\Rightarrow\; \bot}}{\Rightarrow\; \bot} \text{ (Cut)}$$

[8] Here I give only the left-side Contraction rule, as the right-side Contraction rule is not directly involved in the validity curry derivation. But to avoid triviality by rejecting Contraction, we must reject both the left and right rules.

[9] See [37] for an argument that Part A alone is sufficient to cause problems.

We can therefore only avoid triviality by giving up VP, VD, Cut, or Contraction.

Some substructuralists have used this to argue that we should give up Cut or Contraction, moving to a substructural logic. For example, Ripley argues that substructural logics provide a uniform solution to the validity curry and other paradoxes, while other solutions cannot:

"The nonsubstructuralist, then, must modify the V rules, using yet another finger to plug yet another hole in the crumbling dike. After all, no negation or implication occurs in the validity curry argument, so the nonsubstructuralist's tweaks to negation and implication are beside the point here. The substructuralist, on the other hand, has already addressed this problem, since it too depends on both contraction and cut to cause trouble." [32] (pp.308)

Similarly, Beall and Murzi argue that giving up VP or VD is tantamount to giving up on talking about validity:

"One avenue of reply... is to reject one of VP and VD, and concede that we don't have the resources to talk about validity. This line of response—one of "silence", as it is sometimes called—is no more attractive in the case of validity than it is in the case of truth. On the face of it, we *do* talk about validity; and we should seek to account for this phenomenon, rather than deny the data, or deem it incoherent." [4] (pp. 158).

If this is right, then that leaves only Cut or Contraction as potential culprits, thus leaving substructural logics as the only game in town. Validity is an important concept; denying the coherence of validity amounts to denying the coherence of talk about what *follows from* what. I take denying the very coherence of validity to be untenable. So if we must reject one of Cut, Contraction, VP, or VD in order to solve the validity curry paradox, and Beall and Murzi are right that rejecting VP or VD amounts to giving up on talk of validity, then our only real options are to reject Cut or Contraction, thus moving to a substructural logic.

I am skeptical that Beall and Murzi are right here. While VP and VD are *prima facie* very appealing rules, I see no reason to think that denying either rule mounts to a concession that we don't have the resources to talk about validity. But nevertheless, it is true that the validity curry paradox poses a difficult problem for non-substructural approaches. However, as we will see, the validity curry phenomenon can be replicated at higher levels of inference, resulting in an infinite hierarchy of validity curry paradoxes.

As we will see, none of the four rules involved in the paradox—Cut, Contraction, VP, and VD—provide a particular appealing route to solving the full hierarchy of validity curry paradoxes.

4.4 A Metainferential Validity Curry Paradox

Let $Val_2(x, y)$ be a predicate that expresses 2-inference (i.e. metainferential) validity. $Val_2(x, y)$ takes as arguments the names of 1-inferences.[10] For example, the formula $Val_2(\langle \Gamma \Rightarrow_1 \Delta \rangle, \langle \Sigma \Rightarrow_1 \Pi \rangle)$ is the sentence saying that the metainference

$$\frac{\Gamma \Rightarrow_1 \Delta}{\Sigma \Rightarrow_1 \Delta}$$

is valid.

Like the Val_1 predicate, our Val_2 predicate will be governed by two analogous rules. These rules are *metameta*inferences, or 3-inferences:

VP_2:

$$\frac{\{\Gamma \Rightarrow_1 \Delta\} \Rightarrow_2 \{\Sigma \Rightarrow_1 \Pi\}}{\Rightarrow_2 \{\Rightarrow_1 Val_2(\langle \Gamma \Rightarrow_1 \Delta \rangle, \langle \Sigma \Rightarrow_1 \Pi \rangle)\}}$$

VD_2:

$$\frac{}{\{\Gamma \Rightarrow_1 \Delta, \Rightarrow_1 Val_2(\langle \Gamma \Rightarrow_1 \Delta \rangle, \langle \Sigma \Rightarrow_1 \Pi \rangle)\} \Rightarrow_2 \{\Sigma \Rightarrow_1 \Pi\}}$$

[10] As mentioned in section 4.2, I will assume that, in addition to names for finite multisets, our language also has names for ordered pairs of finite multisets, i.e. inferences. If one wants to avoid this extra machinery, one may equivalently take Val_2 to be a four-place predicate, taking the names of four multisets as arguments.

To put the rules in a more familiar notation, VP_2 says that if:

$$\frac{\Gamma \Rightarrow_1 \Delta}{\Sigma \Rightarrow_1 \Pi}$$

is valid, then so is:

$$\frac{}{\Rightarrow_1 Val_2(\langle \Gamma, \Delta \rangle, \langle \Sigma, \Pi \rangle)}$$

VD_2 says that, for all finite multisets Γ, Δ, Σ, and Π, the following metainference is valid:

$$\frac{\Gamma \Rightarrow_1 \Delta, \Rightarrow_1 Val_2(\langle \Gamma, \Delta \rangle, \langle \Sigma, \Pi \rangle)}{\Sigma \Rightarrow_1 \Pi}$$

Here $\Rightarrow_1 Val_2(\langle \Gamma, \Delta \rangle, \langle \Sigma, \Pi \rangle)$ is 1-inference that has the empty multiset as its premises, and $Val_2(\langle \Gamma, \Delta \rangle, \langle \Sigma, \Pi \rangle)$ as its conclusion.

These rules are analogous to the original VP and VD, and they appear to be just as difficult to give up. However, VP_2 and VD_2 lead to triviality if metainferences obey the 2-inference analogues of Cut and Contraction. I'll call these 2-Cut and 2-Contraction. Written in $\Rightarrow_2$ notation, they are:

$$\frac{\Gamma \bigcup \{A \Rightarrow_1 B\} \Rightarrow_2 \Delta \qquad \Gamma' \Rightarrow_2 \{A \Rightarrow_1 B\} \bigcup \Delta'}{\Gamma, \Gamma' \Rightarrow_2 \Delta, \Delta'} \text{ 2-Cut}$$

$$\frac{\Gamma \bigcup \{A \Rightarrow_1 B, A \Rightarrow_1 B\} \Rightarrow_2 \Delta}{\Gamma \bigcup \{A \Rightarrow_1 B\} \Rightarrow_2 \Delta} \text{ 2-Contraction}$$

For our current purposes, we do not need to look at these rules in full generality. In the simplest case, 2-Cut is the metametainference (3-inference) from

$$\frac{}{A \Rightarrow_1 B}$$

and

$$\frac{A \Rightarrow_1 B}{C \Rightarrow_1 D}$$

to

$$\frac{}{C \Rightarrow_1 D}$$

Whereas in the simplest case, 2-Contraction is the metametainference (3-inference) from

$$\frac{A \Rightarrow_1 B \qquad A \Rightarrow_1 B}{C \Rightarrow_1 D}$$

to

$$\frac{A \Rightarrow_1 B}{C \Rightarrow_1 D}$$

2-Cut, 2-Contraction, VP_2 and VD_2 lead to triviality, even without Cut and Contraction at the lower levels. We can see this by looking at a metainferential version of the validity curry paradox.

The 2-inference version of the validity curry paradox is derived analogously to the original validity curry paradox. It too can be divided into a Part A and a Part B.

Let π_2 be the sentence equivalent to $Val_2(\langle \Rightarrow_1 \pi_2 \rangle, \langle \Rightarrow_1 \bot \rangle)$. Intuitively, this sentence says that the metainference from $\Rightarrow_1 \pi_2$ to $\Rightarrow_1 \bot$ is valid, where $\Rightarrow_1 \pi_2$ is the 1-inference with no premises and π_2 as its conclusion.

Part A uses VD_2 and 2-Contraction to derive $\{\Rightarrow_1 \pi_2\} \Rightarrow_2 \{\Rightarrow_1 \bot\}$:

$$\cfrac{\cfrac{\cfrac{}{\{\Rightarrow_1 \pi_2, \Rightarrow_1 Val_2(\langle \Rightarrow_1 \pi_2 \rangle, \langle \Rightarrow_1 \bot \rangle)\} \Rightarrow_2 \{\Rightarrow_1 \bot\}}\,(VD_2)}{\{\Rightarrow_1 \pi_2, \Rightarrow_1 \pi_2\} \Rightarrow_2 \{\Rightarrow_1 \bot\}}\,(\text{def. } \pi_2)}{\{\Rightarrow_1 \pi_2\} \Rightarrow_2 \{\Rightarrow_1 \bot\}}\,(\text{2-Cont})$$

So from VD_2 and 2-Contraction alone, it follows that

$$\frac{\Rightarrow_1 \ \pi_2}{\Rightarrow_1 \ \bot}$$

is a valid metainference.

We can then use this in Part B, to derive the 2-sequent $\Rightarrow_2 \{\Rightarrow_1 \bot\}$ using VP_2 and 2-Cut.

Part B:

$$\dfrac{\dfrac{\dfrac{\dfrac{[\text{Part A}]}{\{\Rightarrow_1 \pi_2\} \Rightarrow_2 \{\Rightarrow_1 \bot\}}}{\Rightarrow_2 \{\Rightarrow_1 Val_2(\langle \Rightarrow_1 \pi_2\rangle, \langle \Rightarrow_1 \bot\rangle)\}}\,{\scriptstyle (VP_2)}}{\Rightarrow_2 \{\Rightarrow_1 \pi_2\}}\,{\scriptstyle (\text{def. } \pi_2)} \qquad \dfrac{[\text{Part A}]}{\{\Rightarrow_1 \pi_2\} \Rightarrow_2 \{\Rightarrow_1 \bot\}}}{\Rightarrow_2 \{\Rightarrow_1 \bot\}}\,{\scriptstyle (\text{2-Cut})}$$

Thus $\Rightarrow_2 \Rightarrow_1 \bot$ is a provably valid 2-inference, regardless of what $\bot$ might be. This is a problem. I assume that our metainferences are *sound*, in the sense that $\Rightarrow_2 \{\Gamma \Rightarrow_1 \Delta\}$ entails $\Gamma \Rightarrow_1 \Delta$. If so, then in the presence of weakening, we will have $\Gamma \Rightarrow_1 \Delta$ for any Γ and any (nonempty) Δ. Therefore our logic is trivial, *even if the logic is noncontractive or nontransitive in the usual sense*. The paradox does not involve Cut or Contraction at the first inferential level, and so cannot be solved by the same substructural means as the original validity curry paradox.

But the problems do not stop there. As we will see in the next section, this phenomenon reoccurs at every level of meta^{n}inference. At each level, the paradoxical derivation does not involve Cut or Contraction of any lower level.

4.5 An Infinite Hierarchy of Validity Curry Paradoxes

This metainferential version of the validity curry paradox can be recreated at every inferential level, ad infinitum. Here is the basic strategy:

Given a sentence ϕ, let $\phi^0 := \phi$, let $\phi^1 := \Rightarrow_1 \phi$, let $\phi^2 := \Rightarrow_2 \{\Rightarrow_1 \{\phi\}\}$, and so on such that $\phi^n := \Rightarrow_n \phi^{n-1}$. Note that we have names for each of these in our language; let $\langle \phi^n \rangle$ be the name for ϕ^n.

For each level of inference $n > 0$, let $Val_n(x, y)$ be a predicate meant to represent the validity relation for n-inferences, which takes names for (n-1)-inferences as arguments. Each $Val_n(x, y)$ will obey two rules analogous to (VP) and (VD), which I'll call (VP$_n$) and (VD$_n$). Letting Γ and Δ be (finite multisets of) (n-1)-inferences, these rules are:

VP_n:

$$\frac{\Gamma \Rightarrow_n \Delta}{\Rightarrow_n Val_n(\langle\Gamma\rangle, \langle\Delta\rangle)^{n-1}} \text{VP}_n$$

VD_n:

$$\frac{}{\Gamma, Val_n(\langle\Gamma\rangle, \langle\Delta\rangle)^{n-1} \Rightarrow_n \Delta} \text{VD}_n$$

where $Val_n(\langle\Gamma\rangle, \langle\Delta\rangle)^{n-1}$ is the (n-1)-inference $\Rightarrow_{n-1}\Rightarrow_{n-2} \ldots \Rightarrow_1 Val_n(\langle\Gamma\rangle, \langle\Delta\rangle)$.

For each level n, we introduce a self-referential n-validity curry sentence. Let π_n be a sentence equivalent to $Val_n(\langle(\pi_n)^{n-1}\rangle, \langle\bot^{n-1}\rangle)$

As before, we can divide the derivation into a Part A and a Part B:

Part A uses n-Contraction and VD$_n$ to derive $(\pi_n)^{n-1} \Rightarrow_n \bot^{n-1}$

$$\cfrac{\cfrac{\cfrac{}{\{(\pi_n)^{n-1}, Val_n(\langle(\pi_n)^{n-1}\rangle, \langle\bot^{n-1}\rangle)^{n-1}\} \Rightarrow_n \{\bot^{n-1}\}} \text{VD}_n}{\{(\pi_n)^{n-1}, (\pi_n)^{n-1}\} \Rightarrow_n \{\bot^{n-1}\}} \text{def. of } \pi_n}{\{(\pi_n)^{n-1}\} \Rightarrow_n \{\bot^{n-1}\}} n\text{-Contraction}$$

Part B uses n-Cut and VP$_n$ to derive $\bot^n$:

$$\cfrac{\cfrac{\cfrac{\cfrac{[\text{Part A}]}{\{(\pi_n)^{n-1}\} \Rightarrow_n \{\bot^{n-1}\}}}{\Rightarrow_n \{Val_n(\langle(\pi_n)^{n-1}\rangle, \langle\bot^{n-1}\rangle)^{n-1}\}} (\text{VP}_n)}{\Rightarrow_n \{\pi_n^{n-1}\}} (\text{def. of } \pi_n) \quad \cfrac{[\text{Part A}]}{\{\pi_n^{n-1}\} \Rightarrow_n \{\bot^{n-1}\}}}{\Rightarrow_n \{\bot^{n-1}\}} (n\text{-Cut})$$

We can therefore prove $\Rightarrow_n \ldots \Rightarrow_1 \bot$, regardless of what $\bot$ is. So if we accept all of n-Cut, n-Contraction, VP_n and VD_n at *any* inferential level n, the logic will be trivial. This derivation schema gives us an infinite metainferential hierarchy of validity curry paradoxes. On this schema, setting $n = 1$ gives us the original validity curry paradox, while $n = 2$ gives us the metainferential validity curry paradox from section 4.4. Higher ns will give us higher order validity curries.

Notice that the derivation for level n does not involve k-Cut or k-Contraction for any level $k < n$; nor does it involve (VP_k) or (VD_k) for any $k < n$. At any level n, the paradox cannot be solved by rejecting the structural or operation rules of lower levels.

4.6 The Principle of Uniform Solution

Substructural approaches aim to provide a uniform solution to semantic paradoxes, including paradoxes of validity like the validity curry paradox. Defenders of these approaches argue that they provide a uniform solution, because they block every paradoxical argument at the exact same step. Other nonclassical approaches do not provide a uniform solution, because they must block negation rules for the liar, conditional rules for the curry, and so on. See, for example, Ripley [34]:

> "The real problem with [nonsubstructural approaches]... is this: the nonsubstructuralist deals with the paradoxes piecemeal, missing the general features that are allowing them to arise in the first place. But paradox runs deeper than any particular vocabulary. Tinkering with negation or conditional rules might prevent paradoxes involving negations and conditionals from arising, but it doesn't come to grips with the general phenomenon."

The metainferential hierarchy of validity curries suggests that providing a uniform solution is not *quite* as simple as blocking *one* rule. Simply blocking Cut or Contraction at the first level of inference will not block the metainferential paradoxes of higher levels. However, it is not a difficult leap to suppose that those who give up Cut at the first level may also be willing to give up Cut at higher levels; similarly for those who reject Contraction. If Contraction is problematic at one level, it stands to reason that it might be problematic at any level. Similarly for Cut.

In practice, defenders of substructural logics vary on this point. Ripley [31] argues against applying the nontransitive ST approach to metainferences, opting to keep metainferences transitive. Meanwhile, Weber [40] and Wansing and Priest [39] defend noncontractive logics that do not validate metainferential Contraction. But regardless of what particular substructuralists have so far defended, one would need an awfully strict notion of uniformity to deny that rejecting either Cut or Contraction at every level was a uniform solution. I take it, then, that giving up either VP, VD, Cut, or Contraction at every level suffices to provide a uniform solution.

But therein lies the difficulty. In the following sections, I will argue that none of the four rules can easily be given up at every level, even for those who are already comfortable giving up one of the rules at level 0. The metainferential hierarchy of validity curry paradoxes therefore creates problems even for logics that have already solved the original validity curry paradox.

Furthermore, I take it that we do in fact need to provide a uniform solution to all of the validity curry paradoxes in the hierarchy. It seems clear that these validity curry paradoxes are essentially the same problem reoccurring at different levels of inference. Surely, the validity curry paradoxes of various metainferential levels are all *the same kind of paradox*, in the sense of Priest's Principle of Uniform Solution: "same kind of paradox, same kind of solution" [26]. These paradoxes all have the exact same structure, use analogous rules, and they all involve validity. As such, these paradoxes require a uniform solution: if Cut is the problem at level 1, then Cut must be the problem "all the way up". It would seem fairly ad hoc to solve the n^{th} validity curry by rejecting Cut_n, but solve the $n+1^{th}$ validity curry by rejecting VP_n. If one is going to solve the full metainferential hierarchy of validity curry paradoxes, then one must reject either VP_n, VD_n, Cut_n, or Contraction_n at every level n.

Therefore if the substructuralist wants to provide a uniform solution to the validity curry paradoxes, they must reject either Cut or Contraction for *every* level of inference.

Similarly, if the non-substructuralist is going to give a uniform solution to the validity curry paradoxes, without rejecting either Cut or Contraction (at any level), it appears that they will have to give up VP at every level, or give up VD at every level. If detachment is the problem at one level, surely it must be the problem at all levels.

However, none of these options are particularly appealing. In the next four sections, I will argue that giving up any of these rules at every metainferential level is more problematic than giving up any of the rules at one level would be. I then suggest that we may need to rethink what it means to give a uniform solution to these paradoxes.

4.7 Giving up Cut at every level

Defenders of the nontransitive approach solve the validity curry paradox and other paradoxes by rejecting Cut. In the case of the validity curry paradox, they avoid triviality by accepting both $\vdash \pi$ and $\pi \vdash \bot$ while rejecting $\vdash \bot$.

At least some defenders of the nontransitive approach have resisted the move to nontransitive metainferences. Ripley argues that the defender of ST as a useful consequence relation "has thereby taken no commitments at all" regarding metainferential consequence [31]. Ripley argues that one can accept the valid 1-inferences of ST as the right set of valid 1-inferences, while still remaining agnostic between different notions of n-validity at levels $n > 1$. In particular, Ripley argues that the defender of ST is *not* obligated to move to the logic that is nontransitive at every inferential level.

However, the metainferential validity curry paradoxes pose a problem for Ripley's position. Ripley and other defenders of ST frequently appeal to the fact that ST provides a uniform solution to semantic paradoxes: ST solves all of the semantic paradoxes by rejecting Cut. Even if Ripley is right that in principle one can endorse the validities of one level while remaining agnostic about higher levels, the defender of ST cannot remain agnostic about higher levels *while providing a uniform solution to semantic paradoxes.*

If defenders of the nontransitive approach want to provide a uniform solution to the validity curry paradoxes, then they will need to give up n-Cut at every level n. They simply cannot remain agnostic about metainferences. There are already formal systems that do this available in the literature.[11] Unfortunately, these are already known to have serious problems.

[1] and [24] introduce a logic which applies the ST nontransitive phenomenon at every metainferential level. If the defender of ST—the most common nontransitive approach to semantic paradoxes in recent years—wants to provide a uniform solution to all of the paradoxes in the validity curry hierarchy, this is the logic that they will need to use.

Unfortunately, this logic is already known to have serious problems. Scambler [35] discusses this logic, and argues convincingly that it puts us in a position much like that of Lewis Carroll's Tortoise: namely, it calls inferences *valid* but does not allow us to *use* those inferences. The set of valid 1-inferences is not closed under the set of valid 2-inferences, the set of valid 2-inferences is not closed under the valid 3-inferences, and so on.

This means that we can start a sequent calculus proof with all valid premises, make nothing but valid metainferences, and still end up with an invalid conclusion.

This is already true of ST: if we start with true premises, and run an argument with all

[11] See [1], [2] and [24] for the technical machinery to construct logics that are not closed under Cut at any metainferential level. These build a hierarchy of nontransitive logics which have the same validities of classical logic at every metainferential level, but which are not closed under Cut at any inferential level. Fitting, in [17], [18], and [19], shows that this can be generalized far beyond classical logic.

valid inferences, then we can still end up with an false conclusion. In ST—with a transparent truth predicate added to the language—the inference from $0 = 0$ to the liar sentence is valid, and the inference from the liar sentence to $0 = 1$ is valid. If we were to actually *make* those inferences, then we could prove $0 = 1$ from $0 = 0$. This is obviously ridiculous.

But, as I argued in "Supervaluations and the Strict-Tolerant Hierarchy", this is not a problem for ST *as long as* the ST 1-inferences are not understood as true *inferences* that we can make. In Ripley's bilateralist interpretation of ST, these "inferences" are really *positions*: they are pairs of assertions and denials. We can use the metainferences in ST to prove which pairs of assertions and denials are incoherent. This is the value of ST: it is not in making inferences from one sentence to another, but in making (meta)inferences from one *position* to another.

The problem with rejecting Cut at every metainferential level is that it leaves us with no true *inferences*. The logic ST is transitive at the metainferential level, and so we can actually *use* the valid metainferences to reason about the "inferences" (positions) at the first level. But if we take our metainferences to be nontransitive *at every level*, then there is no level left at which we can actually *make* the inferences that we're calling valid.

This puts the nontransitive approach into a much worse position than it was in at the first level. The use a logic *as a logic*, we need to be able to string valid inferences together to construct proofs. If defenders of the nontransitive approach aim to provide a uniform solution to the semantic paradoxes by blocking Cut at every inferential level, they will end up with what Scambler calls the "Tortoise logic", and be unable to actually *use* their logic by making valid inferences.

4.8 Giving up Contraction at every level

If defenders of the noncontractive approach want to block all of the validity curry paradoxes in the hierarchy at the same step of the argument, then they will need to reject n-Contraction at every inferential level. This makes the noncontractive approach much less appealing.

Although the noncontractive approach rejects contraction in general, defenders of the approach can take some solace in the fact that there are *forms* of contraction that are still acceptable. For example, [5] point out that Cut allows the noncontractivist to contract on theorems:

$$\frac{\Rightarrow A \qquad \Gamma, A, A \Rightarrow \Delta}{\Gamma, A \Rightarrow \Delta.}\text{Cut}$$

Thus as long as we have $\Rightarrow A$, we can in effect contract on A.

Mares and Paoli [23] make a similar point. They distinguish between *internal* and *external* consequence. Roughly, A is an internal consequence of B if $A \vdash B$ is valid. A is an external consequence of B iff

$$\frac{\vdash A}{\vdash B}$$

is valid. The Contraction rule that noncontractivists reject is contraction for *internal* consequence. In defending their noncontractive system, Mares and Paoli argue that we can tolerate the failure of contraction in *internal* consequence, because we do get some contraction back: *external* consequence contracts. Mares and Paoli distinguish two senses in which we can say A "follows from" Γ. The first sense they call "information extraction". On this sense, A follows from Γ iff "given the rules of the logic at issue, we can extract the information that A from the combined information provided by the sentences in Γ" [23]. The second notion they call "information preservation". On this sense, A follows from Γ iff "Γ yields grounds for asserting A; i.e. whenever we accept Γ we are committed to accepting A" [23].

Mares and Paoli argue that Contraction fails for the first sense, because in some cases you *need* two applications of a premise or axiom in a proof. Sometimes there just is no proof in which the premise is applied only once. Thus, in this sense of “follows from”, there are cases where B follows from A, A but not from A. Thus Mares and Paoli argue that Contraction fails for the first sense of information extraction, even though it does not fail for information preservation. Mares and Paoli take *internal consequence* to be the right consequence for information extraction, and they take *external consequence* to be the right notion of consequence for information preservation.

However, none of this applies if the noncontractivist must reject Contraction at every metainferential level. Contracting on theorems and external consequence contraction both require metainferential contraction. External consequence is just a special case of *internal* metainferential consequence. So if metainferential internal consequence doesn’t contract, then neither does external (1-)consequence. Once we give that up, there is no form of contraction that the noncontractive approach can accept. In Mares and Paoli’s terminology, there is no form of information preservation that the noncontractive approach can accept. The noncontractive approach must reject information preservation as a coherent notion. This problematic, for many reasons. One reason is that information preservation seems to be a perfectly coherent notion, and a very useful one at that. If one’s approach to semantic paradoxes forces us to reject the notion as incoherent, so much the worse for one’s approach to semantic paradoxes. Another, more practical reason this is problematic is that it means that we have no easy way to add non-logical axioms to our system. In the presence of metainferential contraction, we could take A to be an axiom by adding $\Rightarrow_1 A$ to our logic, and that sufficed for us to use A as many times as we liked.

Once we give up 2-Contraction, it doesn’t just matter how many times we use A in a proof; it matters how many times we use $\Rightarrow_1 A$ too. If we want to add non-logical axioms to our logical system, as we often want to do, we have two ways to add them (in a sequent calculus presentation): we can start with our axioms Γ on the left hand side of a sequent, and see what we can prove on the right side. Or, we can add $\Rightarrow_1 \gamma$ for every $\gamma \in \Gamma$ to our sequent calculus, and see what sequents we can prove from them.

The former involves reasoning from the axioms using *internal consequence*, and the latter involves using *external consequence*.

If we have 2-Contraction, then adding $\Rightarrow_1 \gamma$ for all $\gamma \in \Gamma$ allows us to use each γ as many times as we want in derivations; we can always use Cut to remove excess γs from sequents. For example, if we prove $\gamma, \gamma, \Rightarrow_1 \Delta$, we can use Cut twice to get $\Rightarrow_1 \Delta$:

$$\dfrac{\Rightarrow_1 \gamma \qquad \dfrac{\Rightarrow_1 \gamma \qquad \gamma, \gamma \Rightarrow_1 \Delta}{\gamma \Rightarrow_1 \Delta}}{\Rightarrow_1 \Delta}$$

However, if we drop 2-Contraction, this fails. Adding $\Rightarrow_1 \gamma$ doesn't allow us to use γ as many times as we want, because we can no longer use $\Rightarrow_1 \gamma$ as many times as we want to remove excess γs from sequents. Inferences can be "used up" in proofs just like sentences in noncontractive proofs. If we add $\Rightarrow_1 \gamma$ to our logical system, it will matter *how many times we add it*. If $\Gamma, \gamma, \gamma \Rightarrow_1 B$ is valid, and we add only one occurrence of $\Rightarrow_1 \gamma$ to our system, then $\Gamma \Rightarrow_1 B$ may still not be valid. This problem repeats at every metainferential level; for any n and any A, adding $\Rightarrow_n \Rightarrow_{n-1} \ldots \Rightarrow_1 A$ once will give different results than adding it twice. There is no level n that allows us to add an n-theorem to the logic and freely use the theorem more than once.

The problem is that often what we are interested in are the consequence of a collection of assumptions, or a collection of nonlogical axioms, such as those of ZFC. In at least some contexts, we are interested in which we can prove from a sentence or collection of sentences when we have the freedom to use them as many times as we'd like. If we only give up Contraction at the first level, as Mares and Paoli [23], Zardini [42], and Cintula and Paoli [5] do, then we can do this by adding $\Rightarrow A$ to our system for every axiom A that we want. But once we give up Contraction at every metainferential level, there is no way to do this.

What the considerations in this section show is that this can never be done in a noncontractive logic, at least if the logic aims to provide a uniform solution to the validity curry paradoxes. Noncontractive logics do not allow us to introduce nonlogical axioms or other assumptions *with the explicit approval to use them as many times as we'd like.* This means that, in contexts in which we want to use sentences as many times as we'd like, such as when asking about the consequences of ZFC or other mathematical sets of axioms, we cannot use a noncontractive logic. Noncontractive logics are, therefore at best incomplete solutions. Whatever their merits may be, they do not tell us how to solve the paradoxes in contexts where we want to discover consequences of a set of nonlogical axioms or assumptions.

4.9 Giving up VP at every level

At first glance, VP looks like the most obvious step to block:

Validity Proof:

$$\frac{\Gamma \Rightarrow \Delta}{\Rightarrow Val(\langle \Gamma \rangle, \langle \Delta \rangle)} \text{(VP)}$$

As Field points out in [14], this is a bit odd. If $Val(x, y)$ is meant to represent $\Rightarrow$, then validity proof has *one* occurrence of a validity symbol in the premise ($\Rightarrow$), but *two* in the conclusion ($\Rightarrow$ and Val). The conclusion of the metainference VP is therefore, at least on many consequence relations, a strictly stronger claim than the premise. Compare this to the naive truth rules, such as:

$$\frac{A \Rightarrow B}{A \Rightarrow T\langle B \rangle}$$

On most theories of truth, the conclusion here is not *stronger* than the premise. The premise and conclusion are equivalent; that is at least part of why it is so difficult to reject the inference.

But in the case of VP, the conclusion does appear to be something stronger than the premise, which gives us room to reject VP. The conclusion appears to say more than simply that the inference from Γ to Δ is valid; it says that the validity of the inference is a *theorem* in our consequence relation $\Rightarrow$.

It's not at all obvious that we need VP in order to have a naive theory of validity, because it's not obvious that a naive theory of validity demands that we be able to *prove* the relevant Val-sentence for every valid inference. What we really want is to be able to assert that Val-sentence, or for that Val-sentence to be *true*.

Unfortunately, as Shapiro [37] has shown, even this weaker version of VP will lead to paradox. Shapiro introduces a rule that he calls VT:

$$\text{(VT) } Val(\langle A \rangle, \langle B \rangle) \text{ is true iff } A \vdash B.$$

Shapiro shows that replacing VP with the right-to-left direction of still suffices for paradox. Recall that part A of the derivation gives us $\pi \Rightarrow \bot$:

$$\dfrac{\dfrac{\dfrac{}{\pi, Val(\langle \pi \rangle, \langle \bot \rangle) \;\Rightarrow\; \bot}\text{(VD)}}{\pi, \pi \;\Rightarrow\; \bot}\text{(def. of } \pi)}{\pi \;\Rightarrow\; \bot}\text{(Contraction)}$$

Applying Shapiro's VT rule, we get:

$$Val(\langle \pi \rangle, \langle \bot \rangle) \text{ is true}$$

But $Val(\langle \pi \rangle, \langle \bot \rangle)$ is equivalent to π, so this means that we have

$$\pi \text{ is true}$$

We therefore have a valid inference $\pi \vdash \bot$, which has a true premise. If validity is truth preserving, then $\bot$ is true, and we again have triviality.

One potential issue here is that not everyone will accept that validity is truth-preserving.[12] Another issue is that not everyone will accept that theorems $\Rightarrow A$ are *true*.[13]

However, we can formalize the argument without introducing a natural language truth predicate into the mix, by introducing a standard model. Let's assume that we have a standard model, which I will denote by α, that gets all of the (first-level) validity facts right. In other words, $\alpha \Vdash Val(\langle A\rangle, \langle B\rangle)$ iff $A \Rightarrow B$.

Part A of the argument is unchanged, and again gives us $\pi \Rightarrow \bot$. Given our assumption about the standard model, this gives us:

$$\alpha \Vdash Val(\langle \pi\rangle, \langle \bot\rangle)$$

But $Val(\langle \pi\rangle, \langle \bot\rangle)$ is equivalent to π, so this means that we have

$$\alpha \Vdash \pi$$

Given VD, we have $\pi, Val(\langle \pi\rangle, \langle \bot\rangle) \Rightarrow \bot$. We therefore have:

$$\alpha \Vdash \bot$$

Therefore the standard model is trivial. This means that even without VP in its usual form, we cannot have even a single model that gets the validity facts right. This poses a problem. Part of why we want to formally represent validity is to provide a model (in the scientists' sense of the term) of our own, pre-formal notion of *real* validity. If we cannot even have a single model (in the logician's sense) that gets the validity facts right, then we have simply failed to model (in the scientists' sense) real validity.

Defenders of nontransitive and noncontractive approaches will argue that this version of the argument implicitly uses Cut and Contraction, both of which are somewhat hidden by the switch to reasoning about satisfaction at α. Having a model theory with a single

[12]See especially [16] and [15].
[13]See especially [33] and [32].

notion of truth-in-the-model is incompatible with both approaches, and so by reasoning about such a model I am making certain assumptions that neither flavor of substructuralist will accept. I am willing to concede that point; this argument does not directly threaten the noncontractive or nontransitive approaches.

However, this argument and Shapiro's argument *do* threaten any approach that aims to block the validity curry paradoxes by rejecting VP at every level. Rejecting the usual form of VP is not sufficient to block the validity curry paradox at the first level of inference. If we aim to reject VP, we must further deny that there can be any notion of a standard model that gets all of the validity facts right. Furthermore, we must either deny that $A \Rightarrow B$ entails that it is true that the inference from A to B is valid, or we must accept that there are valid inferences with true premises with conclusions that we cannot accept on pain of triviality.

Note that this same reasoning does not apply to the metainferential validity curries. To see why, let's examine the 2-validity curry again. In part A, we derive $\{\Rightarrow_1 \pi_2\} \Rightarrow_2 \{\Rightarrow_1 \bot\}$. Let's suppose that we replace VP with either Shapiro's VT or the standard-model-restricted version of VP, and thus we have either:

$$Val_2(\langle\Rightarrow_1 \pi_2\rangle, \langle\Rightarrow_1 \bot\rangle) \text{ is true}$$

or:

$$\alpha \Vdash Val_2(\langle\Rightarrow_1 \pi_2\rangle, \langle\Rightarrow_1 \bot\rangle)$$

Thus we have the truth of the Val-sentence, or at least truth-in-the-standard-model of the Val-sentence. What we do *not* have is the *theoremhood* of the Val-sentence. Without that, we cannot use $\{\Rightarrow_1 \pi_2\} \Rightarrow_2 \{\Rightarrow_1 \bot\}$ to derive $\bot$ or $\Rightarrow_1 \bot$ or $\Rightarrow_2 \Rightarrow_1 \bot$.

Thus at higher levels of validity, rejecting VP_n in favor of one of these weaker rules merely forces us to say that the Val-sentence is true but not a theorem. Perhaps that is a somewhat awkward position, but it is nowhere near as problematic as the position that rejecting VP at the *first* level would put us in.

However, if we must provide a uniform solution to the validity curry paradoxes by blocking the same step of the argument at every level, then the problems for rejecting VP at level 1 suffice to rule out rejecting VP as an option. We must instead look elsewhere.

4.10 Giving up VD at every level

Giving up VD at first appears more difficult than giving up VP, but the move has its defenders. Field [14] argues that VD fails for at least some notions of validity, while while Hlobil [21] argues that the nontransitive approach should reject VD.

However, these only consider VD at the *first* level. Rejecting VD at higher levels of the validity curry paradox is much harder. Because unlike rejecting VD at the first level, rejecting VD at metainferential levels results in false theorems.

Consider the instance of VD for the 2-validity curry paradox:

$$\{\Rightarrow_1 \pi_2, \Rightarrow_1 Val_2(\langle \Rightarrow_1 \pi_2\rangle, \langle \Rightarrow_1 \bot\rangle)\} \Rightarrow_2 \{\Rightarrow_1 \bot\}$$

As discussed in section 4.2.1, the notion of metainferential validity under discussion here is *global* metainferential validity. So to reject this metainference, we must take the premises of the metainference to be valid, while denying that the conclusion of the metainference is valid. That means that we must take

$$\Rightarrow \pi$$

and

$$\Rightarrow Val_2(\langle \Rightarrow_1 \pi_2\rangle, \langle \Rightarrow_1 \bot\rangle)$$

to be valid, while denying that

$$\Rightarrow \bot$$

is valid. But if $\Rightarrow \pi$ is valid and $\Rightarrow \bot$ is not, then the metainference

$$\frac{\Rightarrow \pi}{\Rightarrow \bot}$$

cannot be valid. And yet,

$$\Rightarrow Val_2(\langle \Rightarrow_1 \pi_2 \rangle, \langle \Rightarrow_1 \bot \rangle)$$

is valid.

This means that we have a false theorem: the metainference from $\Rightarrow \pi$ to $\Rightarrow \bot$ is not valid, yet the sentence *saying* that it's valid is a theorem of our logic.

Some, such as Ripley [34], do not claim that the theorems of their logic are true. Ripley presents a bilateralist view according to which the theorems of ST are simply the sentences which it is incoherent to *deny*; we have no obligation to *accept* them. But this bilateralist response is not satisfying, particularly when the theorem is something we absolutely *can* deny. If we reject this case of (VD), then it is a simple and straightforward fact about the logic that the metainference from $\Rightarrow \pi$ to $\Rightarrow \bot$ is not valid. Ripley's bilateralist picture (on which denying A is equivalent to asserting $\neg A$) then tells us that it is incoherent to assert basic facts about our logic.

That seems unreasonable. Even if we are not committed to accept the theorems of our logic—which, for the record, I think we are—we are certainly free to coherently assert demonstrable facts about our logic.

In short, giving up VD in the metainferential case seems to be an untenable solution. As a result, even if giving up VD appears to be a good option in the original validity curry paradox, rejecting VD "all the way up" is not a tenable uniform solution to the validity curry paradoxes at all levels.

4.11 Rethinking Uniformity

On one notion of uniform solution, a uniform solution to the validity curry paradoxes would have to involve blocking the same step of the argument in every case. This is the notion of uniformity on which substructural logics allegedly provide a uniform solution to semantic paradoxes, while other non-classical solutions allegedly do not provide a uniform solution.

However, blocking the same step of the validity curry argument at every level of inference is not at all appealing. Rejecting Cut at every level leaves us with what Scambler calls the "Tortoise Logic", unable to make the inferences that we accept as valid. Rejecting Contraction at every level leaves us with no way to add non-logical axioms and use them freely in arguments. Rejecting VP at every level fails to solve the paradox at the first level, and rejecting VD at every level requires accepting false theorems about metainferential inferences.

This means that we may need to rethink what it would mean to provide a uniform solution to the paradoxes of the validity curry hierarchy. To that end, it is worth noting that the validity curry paradox of the first level involves a different notion of validity than the metainferential validity curry paradoxes. As I noted in 4.2.1, there are multiple notions of metainferential validity that we might use. There are the semantic notions of *local* and *global* validity, and the proof-theoretic notions of *derivable rules* and *admissible rules.* I said that the notion of metainferential validity in play was the *global* notion. But the notion of validity at the *first* level—the notion of validity between (multisets of) sentences—is distinctly *not* a global notion of validity. Recall the various definitions metainferential validity:

Definition 4.11.1. *Local Metainferential Validity*

A metainference

$$\frac{\Gamma \vdash \Delta}{\Sigma \vdash \Pi}$$

is locally valid iff at every model m*, if* $\Gamma \vdash \Delta$ *is satisfied at* m*, then* $\Sigma \vdash \Pi$ *is satisfied at* m*.*

Definition 4.11.2. *Global Metainferential Validity*

A metainference

$$\frac{\Gamma \vdash \Delta}{\Sigma \vdash \Pi}$$

is globally valid iff if (materially) $\Gamma \vdash \Delta$ *is valid (satisfied at all models), then* $\Sigma \vdash \Pi$ *is valid.*

These are both model-theoretic notions of metainferential validity. But there are also two closely related proof-theoretic notions that we might be interested in:

Definition 4.11.3. *Derivable Metainferential Rule*

A metainference

$$\frac{\Gamma \vdash \Delta}{\Sigma \vdash \Pi}$$

is a derivable rule iff there exists a proof from $\Gamma \vdash \Delta$ *to* $\Sigma \vdash \Pi$*.*

Definition 4.11.4. *Admissible Metainferential Rule*

A metainference

$$\frac{\Gamma \vdash \Delta}{\Sigma \vdash \Pi}$$

is an admissible rule iff if (materially) there exists a proof of $\Gamma \vdash \Delta$*, then there exists a proof of* $\Sigma \vdash \Pi$*.*

We could easily consider analogous notions for regular, non-meta validity. Instead of using the notion of *satisfaction*-at-a-model, we simply use the notion of *truth*-in-a-model.

Definition 4.11.5. *Local 1-Validity*

An inference

$$\Gamma \vdash \Delta$$

is locally valid iff at every model m*, if every* $\gamma \in \Gamma$ *is true at* m*, then some* $\delta \in \Delta$ *is true at* m*.*

Definition 4.11.6. *Global 1-Validity*

An inference

$$\Gamma \vdash \Delta$$

is globally valid iff if (materially) every $\gamma \in \Gamma$ *is valid (true in all models), then some* $\delta \in \Delta$ *is valid.*

Definition 4.11.7. *Derivable 1-Rule*

An inference

$$\Gamma \vdash \Delta$$

is a derivable rule iff there exists a proof from Γ *to some* $\delta \in \Delta$*.*

Definition 4.11.8. *Admissible 1-Rule*

An inference

$$\Gamma \vdash \Delta$$

is an admissible rule iff if (materially) there exists a proof of γ *for all* $\gamma \in \Gamma$*, then there exists a proof of* δ *for some* $\delta \in \Delta$*.*

All four of these notions of 1-validity are perfectly coherent notions. However, it's clear that ordinarily, the notions of 1-validity that we care about are the *local* and *derivable* versions. These are the model-theoretic and proof-theoretic notions of validity that we are most familiar with *for 1-inferences*.

As such, there is a distinction to be drawn between the validity curry paradox at the first level of inference, and the validity curry paradoxes at the higher levels. As presented above, the first level uses a *local* or *derivable* notion of validity, while the metainferential versions use a *global* or *admissible* notion of validity.

As such, although the paradoxical arguments of every level have the same structure, and are *syntactically* analogous, there is a potentially significant difference between the notions of validity involved.

I propose that a uniform solution to the validity curry paradoxes does not require that the same step of the argument be blocked for every notion of validity involved. In particular, I think that we are free to block VP for the *global* notions of validity (at any level), and free to block VD for the *local* notions of validity (at every level).

Note that the definitions for global and admissible validity are material conditionals. This means that whatever solution to the curry paradox we already accept for the material conditional should also apply to the global and admissible versions of the validity curry (at any level). For the paracomplete approach, this means rejecting Conditional Proof (CP), which would translate here to rejecting (VP). Whatever justifications the paracomplete theorist has offered for the failure of (CP) for the material conditional, those should transfer automatically to global and admissible validity curries.

Meanwhile, the local and derivable definitions above are distinctly *not* material conditionals. The definition of a derivable rule is an existential claim: *there exists* a proof. The definition of the local notion is a sort of *modal* claim: *at all models*, a certain condition holds. As such, solutions for the material conditional curry will not automatically carry over to the local and derivable versions of validity curry.

For the local and derivable notions, I believe that the problem in the validity curry argument is VD. Consider the following natural deduction presentation of the validity curry argument, as it applies to 1-validity:

1	$\quad\pi$	Ass.
2	$\quad Val(\langle\pi\rangle, \langle\bot\rangle)$	1 (def of π)
3	$\quad\bot$	VD 1, 2
4	$Val(\langle\pi\rangle, \langle\bot\rangle)$	VP 1-3
5	π	4 (def of π)
6	$\bot$	VD 4,5

Since π is equivalent to (or just *is*) $Val(\langle\pi\rangle, \langle\bot\rangle)$, we have effectively assumed that the inference from π to $\bot$ is valid. Using VD inside that assumption is, in effect, *making* the inference from π to $\bot$. But assuming hypothetically that the inference is valid does not justify making the inference. We should only make inferences that are *actually* valid, not inferences that we have hypothetically *assumed* to be valid. If the inference from π to $\bot$ is not valid, then (since local validity is essentially a claim about logical necessity) the assumption that it *is* valid amounts to a counterlogical assumption, and the use of VD under that assumption amounts to the use of an invalid inference. We should therefore reject VD, because in at least some cases it licenses the use of invalid inferences in proofs.

VD looks even worse on the derivable notion of validity. On this notion, the assumption of π is effectively an assumption that a certain proof exists. The application of VD under the assumption allows us to infer $\bot$ from π *as if* there were such a proof, but there may be no such proof. We have only assumed that there is a proof, and then inferred *as if there* *were one*. This is illegitimate. We can infer $\perp$ from π when there *is* a proof, not when we have *assumed* that there is a proof.

Note that this reasoning against VD does not apply to the global or admissible notions of validity: on these notions, the application of VD is just an application of modus ponens (MP) for the material conditional. The problem, at least for the paracomplete approach, is the application of VP, which is in actuality an application of Conditional Proof.

I do not mean to propose that this explanation of the phenomenon will be acceptable to every non-classical approach; in particular, paraconsistent approaches that reject MP and accept CP for the material conditional will reject the approach that I've outlined. The important point is rather that the non-substructural approaches can give a uniform solution to the validity curry paradoxes, and can do so *without* having to block the same step of the argument in every version.

4.12 Conclusion

Substructural approaches to semantic paradoxes face a problem. They aim to provide a uniform solution to all of the paradoxes: give up either Cut or Contraction, and every paradox will fall. However, the situation is not so simple. There are metainferential validity curry paradoxes, none of which use Cut or Contraction of any lower level of inference. Giving up Cut or Contraction of the first level of inference, as the substructural approaches have so far done, will not solve these paradoxes.

To solve the metainferential validity curry paradoxes, sub nontransitive and noncontractive approaches must reject the metainferential versions of Cut or Contraction at *every* inferential level. This makes the nontransitive and noncontractive approaches much less appealing. Giving up Cut or Contraction at every level of inference leaves us unable to introduce non-logical axioms and discover their consequences in the way that we often to do when we use a logic.

Furthermore, the notion of uniform solution that defenders of substructural approaches often assume is a problematic one. Substructural approaches are often defended by appealing to the fact that they block every paradox at the same step, while other non-classical approaches must reject different rules to prevent different paradoxes. If I am right that the substructural approaches do not solve the paradoxes in full generality, and do not solve the paradoxes in contexts in which we want to discover the consequences of non-logical axioms, then this notion of uniformity used to defend substructural logics is simply mistaken. Neither VP nor VD can tenably be rejected to solve the validity curry paradoxes of *every* level. As such, providing a uniform solution to the validity curry paradoxes requires a more fine-grained notion of uniformity. Different notions of validity will result in different versions of the validity curry paradox, and each of the validity rules may apply to some notions of validity but not to all notions of validity.

I close with a look to the future: this paper has introduced metainferential validity curry paradoxes. In [25], Priest introduces metainferential versions of the liar and conditional curry paradox. Metainferential paradoxes are a new and exciting area, and there is much that is still left to be explored. For example, to the best of my knowledge, the possibility of metainferential set-theoretic paradoxes, and metainferential vagueness paradoxes, have yet to be investigated. What consequences those investigations may have for substructural and other non-classical approaches remains to be seen.

Bibliography

[1] Barrio, Eduardo; Pailos, Federico; Szmuc, Damian (2020) “A Hierarchy of Classical and Paraconsistent Logics” *Journal of Philosophical Logic*, 49: 93-120.

[2] Barrio, Eduardo; Pailos, Federico; Szmuc, Damian (2019) "(Meta)inferential Levels of Entailment Beyond the Tarskian Paradigm" *Synthese*, 198(22): 5265-5289.

[3] Barrio, Eduardo; Rosenblatt, Lucas; and Tajer, Diego (2016) "Capturing Naive Validity in the Cut-Free Approach" *Synthese* https://doi.org/10.1007/s11229-016-1199-5.

[4] Beall, JC and Murzi, Julien (2013) "Two Flavors of Curry's Paradox" *Journal of Philosophy*, 110 (3): 143-165.

[5] Cintula, Petr and Paoli, Francesco (2016) "Is Multiset Consequence Trivial?" *Synthese*. https://doi-org.ezproxy.gc.cuny.edu/10.1007/s11229-016-1209-7.

[6] Cobreros, Pablo; Egré, Paul; Ripley, David; van Rooij, Robert (2013) "Reaching Transparent Truth" *Mind*, 122(488): 841-866.

[7] Cobreros, P., Egré, P., Ripley, D., van Rooij, R. (2012a). "Tolerance and mixed consequence in the S'valuationist setting." *Studia Logica*, 100(4), 855–877.

[8] Cobreros, Pablo; Egré, Paul; Ripley, David; van Rooij, Robert (2012b) "Tolerant, Classical, Strict" *Journal of Philosophical Logic*, 41 (2): 347-385.

[9] Da Ré, Bruno, Szmuc, Damian, and Teijeiro, Paula (2021) "Derivability and Metainferential Validity" *Journal of Philosophical Logic* https://doi.org/10.1007/s10992-021-09619-3.

[10] Došen, K. (1985a) "Sequent-Systems for Modal Logic" *Journal of Symbolic Logic*, 50: 149-159.

[11] Došen, K. (1986) "Higher-Level Sequent Systems for Intuitionistic Modal Logic" *Publications de l'Institut Mathématique*, 39(53): 3-12.

[12] Došen, K. (1988) "Sequent Systems and Groupoid Models I" *Studia Logica*, 47: 353-389.

[13] Došen, K. (1992) "Modal Logic as Metalogic" *Journal of Logic, Language, and Information* 1: 173-201.

[14] Field, Hartry (2017) "Disarming a Paradox of Validity" *Notre Dame Journal of Formal Logic*, 58(1): 1-19.

[15] Field, Hartry (2015) "What is Logical Validity?" *Foundations of Logical Consequence* ed. Colin Caret, Ole Hjortland (Oxford University Press).

[16] Field, Hartry (2008) *Saving Truth from Paradox* (Oxford University Press).

[17] Fitting, Melvin (2021a) "The Strict/Tolerant Family Continued: Quantifiers and Modalities" *The Australasian Journal of Logic*, 18(6): 616-644.

[18] Fitting, Melvin (2021b) "A Family of Strict/Tolerant Logics" *Journal of Philosophical Logic*, 50: 363–394.

[19] Fitting, Melvin (2021c) "The Strict/Tolerant Idea and Bilattices" in *For Arnon Avron*, ed. Ofer Arieli et al., Outstanding Contributions to Logic.

[20] French, Rohan (2016) "Structural Reflexivity and the Paradoxes of Self-Reference" *Ergo*, 3 (5): 113-131.

[21] Hlobil, Ulf (2019) "Faithfulness for Naive Validity" *Synthese*, 196 (11): 4759-4774.

[22] Hlobil, Ulf (2018) "The Cut-Free Approach and the Admissibility-Curry" *Thought: A Journal of Philosophy*, 7(1):40-48.

[23] Mares, Edwin and Paoli, Francesco (2014) "Logical Consequence and the Paradoxes" *Journal of Philosophical Logic*, 43 (2-3), 439-469.

[24] Pailos, F. (2020). "A Fully Classical Truth Theory Characterized by Substructural Means" *The Review of Symbolic Logic*, 13 (2): 249-268.

[25] Priest, Graham (manuscript) "Substructural Solutions to the Semantic Paradoxes: Dialetheism in Sheep's Clothing?"

[26] Priest, Graham (1994) "The Structure of the Paradoxes of Self-Reference" *Mind*, 103: 25-41.

[27] Rosenblatt, Lucas (2019a) "Non-Contractive Classical Logic" *Notre Dame Journal of Formal Logic*, 60 (4), 559-585.

[28] Rosenblatt, Lucas (2019b) "On Structural Contraction and Why It Fails" *Synthese*. https://doi.org/10.1007/s11229-019-02241-6.

[29] Rosenblatt, Lucas (2017) "Naive validity, Internalization, and Substructural Approaches to Paradox" *Ergo*, 4 (4): 93-120.

[30] Wansing, Heinrich and Priest, Graham (2015) "External Curries" *Journal of Philosophical Logic*, 44: 453-471.

[31] Ripley, David (2021) "One Step is Enough" *Journal of Philosophical Logic*. https://doi.org/10.1007/s10992-021-09615-7.

[32] Ripley, David (2015) "Comparing Substructural Theories of Truth" *Ergo*, 2(13):299-328.

[33] Ripley, David (2013) "Paradoxes and Failures of Cut" *Australasian Journal of Philosophy* 91 (1): 139-164.

[34] Ripley, D. (2012) "Conservatively Extending Classical Logic with Transparent Truth" *Review of Symbolic Logic*, 5 (2): 354-378.

[35] Scambler, C. (2020) "Classical Logic and the Strict Tolerant Hierarchy" *Journal of Philosophical Logic*, 49: 351-370.

[36] Shapiro, Lionel (2015) "Naive Structure, Contraction, and Paradox" *Topoi* 34 (1): 75-87.

[37] Shapiro, Lionel (2013) "Validity Curry Strengthened" *Thought*, 2, 100-107.

[38] Shapiro, Lionel (2011) "Deflating Logical Consequence" *Philosophical Quarterly*, 61 (243), 320-342.

[39] Wansing, Heinrich and Priest, Graham (2015) "External Curries" *Journal of Philosophical Logic*, 44: 453-471.

[40] Weber, Zach (2016) "On Closure and Truth in Substructural Theories of Truth" *Synthese*. https://doi-org.ezproxy.gc.cuny.edu/10.1007/s11229-016-1226-6.

[41] Zardini, Elia (2019) "Instability and Contraction" *Journal of Philosophical Logic*, 48:155-188.

[42] Zardini, Elia (2011). Truth Without Contra(di)ction. *Review of Symbolic Logic*, 4 (4), 498-535.